

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 07:41:54 ; Search time 3.2668 Seconds  
(Without alignments)  
7582.941 Million cell updates/sec

Title: US-09-902-772-1\_COPY\_650\_660  
Perfect score: 11  
Sequence: 1 agagagagagag 11

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	11	100.0	21	AAH49300	V. vinifera aquapo
C 2	11	100.0	21	ABK94127	Endothelin-2 (EDN-
C 3	11	100.0	21	ABK94128	Endothelin-2 (EDN-
C 4	11	100.0	22	AAK75373	CDNA synthesis pri
C 5	11	100.0	22	AAV59955	PCR primer EGR11-6
C 6	11	100.0	23	AAQ39034	Oligonucleotide id
C 7	11	100.0	24	AAK18161	CEBP Tag1B polymer
C 8	11	100.0	25	AAH37422	SNP specific lower
C 9	11	100.0	31	AAQ40991	Mono ADP-ribosyltr

C 10	11	100.0	32	18	AAK91953	PCR primer 1 used
C 11	11	100.0	36	22	AAK88029	Human Trp 1 PCR pr
C 12	11	100.0	40	21	AAK295968	Polynucleotide seq
C 13	11	100.0	43	20	AAK81695	Primer used in a n
C 14	11	100.0	51	21	AAA77247	Human clone c94397
C 15	11	100.0	60	24	ABN40086	Human spliced tran
C 16	11	100.0	63	19	AAV29642	Nucleotide sequenc
C 17	11	100.0	63	22	AAK00974	PCR primer for con
C 18	11	100.0	65	24	ABN58012	Mouse spliced tran
C 19	11	100.0	72	13	AAQ33600	Microsatellite seq
C 20	11	100.0	76	13	AAQ34136	Sequence upstream
C 21	11	100.0	90	24	ABK76766	Bacillus lichenifo
C 22	11	100.0	98	24	AAK563133	Cell death protect
C 23	11	100.0	99	22	ABK48277	Human breast cell
C 24	11	100.0	99	22	ABK46159	Human foetal liver
C 25	11	100.0	99	22	ABK33228	Probe #11694 for g
C 26	11	100.0	99	22	AAK14578	Human brain expres
C 27	11	100.0	99	22	AAK40324	Human bone marrow
C 28	11	100.0	99	22	AAK121082	Probe #11015 for g
C 29	11	100.0	99	22	AAK16343	Probe #15029 used
C 30	11	100.0	99	22	AAK106802	Probe #6793 used t
C 31	11	100.0	99	24	ABK14344	Human genome-deriv
C 32	11	100.0	111	14	AAQ39030	Oligonucleotide id
C 33	11	100.0	111	14	AAQ39031	Oligonucleotide id
C 34	11	100.0	111	22	ABK40333	Probe #18799 for g
C 35	11	100.0	113	21	AAV16149	Microsatellite mar
C 36	11	100.0	113	19	AAK17053	Human secreted pro
C 37	11	100.0	115	14	AAQ39032	Oligonucleotide id
C 38	11	100.0	115	14	AAQ39033	Oligonucleotide id
C 39	11	100.0	117	24	ABK20896	Human genome-deriv
C 40	11	100.0	118	19	AAV16144	Microsatellite mar
C 41	11	100.0	124	22	ABK47414	Human breast cell
C 42	11	100.0	124	22	ABK45302	Human foetal liver
C 43	11	100.0	124	22	ABK32403	Probe #10869 for g
C 44	11	100.0	124	22	AAK13716	Human brain expres
C 45	11	100.0	124	22	AAK39457	Human bone marrow

#### ALIGNMENTS

RESULT 1	
AAH49300/c	
ID	AAH49300 standard; DNA; 21 BP.
XX	
AC	AAH49300;
XX	
DT	28-NOV-2001 (first entry)
XX	
DE	V. vinifera aquaporin probe SEQ ID 12.
XX	
KW	Aquaporin; transmembrane protein; intrinsic protein; water transport;
KW	cell growth; drought tolerance; salt tolerance; probe; ss.
XX	
OS	Vitis vinifera.
XX	
PN	W0200166793-A2.
XX	
PD	13-SEP-2001.
XX	
PF	09-MAR-2001; 2001WO-EP02718.
XX	
PR	09-MAR-2000; 2000DE-1011480.
XX	
PA	(GSFUT-) GSF FORSCHUNGSZENTRUM UNWELT & GESUNDHEIT.
XX	
PI	Baiges Blanco IM, Schaeffner AR, Mas A;
XX	
DR	WPI; 2001-589950/66.
XX	
PT	Nucleic acid for amplifying and quantitatively and qualitatively
PT	detecting aquaporin genes; gene fragments, RNA or expression of an
PT	aquaporin gene from Vitis species

XX Claim 1; Page 28; 38pp; German.  
 XX  
 CC This invention describes a novel nucleic acid for amplifying and  
 CC quantitatively and qualitatively detecting aquaporin genes, RNA or gene  
 CC expression. The nucleic acid is used to qualitatively and/or  
 CC quantitatively detect or amplify aquaporin genes, gene fragments or gene  
 CC expression from *Vitis* species. Aquaporin is a transmembrane protein  
 CC involved in water transport and known to affect cell growth and drought  
 CC and salt tolerance. This sequence represents a probe used in the  
 CC detection of the *Vitis vinifera* aquaporin associated with the intrinsic  
 CC protein family.  
 XX  
 SQ Sequence 21 BP; 2 A; 9 C; 1 G; 9 T; 0 other;  
 Query Match 100.0%; Score 11; DB 22; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGAGAGAGAGAG 11  
 Db 20 AGAGAGAGAGAG 10

RESULT 2  
 ABR94127 ID ABR94127 standard; DNA: 21 BP.  
 AC ABR94127;  
 XX  
 DT 27-AUG-2002 (first entry)  
 DE  
 XX Endothelin-2 (EDN-2) SNP detection PCR primer #23.  
 DE  
 XX Endothelin: EDN; endothelin converting enzyme; ECE; endothelin receptor;  
 KW EDNR; signaling system; cardiovascular disease; coronary heart disease;  
 KW hypertension; atherosclerosis; angiogenesis; fatty acid metabolism;  
 KW diabetes; familial hypercholesterolemia; forensic marker;  
 KW transgenic animal; solid support; cardiovascular regulator; SNP;  
 KW single nucleotide polymorphism; PCR; primer; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200224747-A2.  
 XX  
 PD 28-MAR-2002.  
 XX  
 PF 31-AUG-2001; 2001MO-EP10087.  
 XX  
 PR 19-SEP-2000; 2000EP-0120123.  
 XX  
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX  
 PI Brinkmann U, Hoffmeyer S;  
 DR WPI: 2002-435060/46.  
 XX  
 XX Novel polynucleotide of the endothelin/endothelin converting  
 PT enzyme/receptors of endothelin and endothelin converting enzyme  
 PT signaling system associated with cardiovascular disease, useful for  
 PT treating the disease  
 XX  
 Claim 1; Page 57; 190pp; English.

The invention describes a polynucleotide (I) of the endothelin  
 CC (EDN)/endothelin converting enzyme (ECE)/receptors of EDN and ECE (EDNR)  
 CC signaling system which is associated with a cardiovascular disease. (I),  
 CC the gene encoding EDN, ECE or EDNR (II) or a vector (III) expressing (I)  
 CC or (II) is useful for producing cells capable of expressing a molecular  
 CC variant polypeptide which is associated with a cardiovascular disease.  
 CC (II), (III), the EDN, ECE or EDNR polypeptide, or a cell expressing  
 CC a molecular variant gene comprising (I) is useful for identifying and  
 CC obtaining a pro-drug or drug capable of modulating the activity of a

CC molecular variant of a polypeptide of the EDN/EDNR/ECE signaling system  
 CC or its gene product, or for identifying and obtaining an inhibitor of  
 CC the activity of a molecular variant of a polypeptide of the EDN/EDNR/ECE  
 CC signaling system or its gene product. The isolated proteins and  
 CC polynucleotides encoding them are useful for preparation of a  
 CC pharmaceutical composition for treating a cardiovascular disease such as  
 CC coronary heart disease, hypertension, atherosclerosis, or related to  
 CC abnormal angiogenesis or fatty acid metabolism e.g. diabetes and familial  
 CC hypercholesterolemia. The gene or a polynucleotide fragment of the  
 CC EDN/ECE/EDNR signaling system are useful as forensic markers, for  
 CC creating a transgenic animal and in creation of a solid support  
 CC comprising polynucleotides, genes, vectors, polypeptides, antibodies or  
 CC host cells of the invention. This sequence represents a PCR primer used  
 CC to identify single nucleotide polymorphisms in DNA encoding  
 CC cardiovascular regulator proteins of the EDN/ECE/EDNR signaling pathway.  
 XX  
 SQ Sequence 21 BP; 7 A; 4 C; 10 G; 0 U; 0 other;  
 Query Match 100.0%; Score 11; DB 24; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGAGAGAGAGAG 11  
 Db 11 AGAGAGAGAGAG 21

RESULT 3  
 ABR94128/C ID ABR94128 standard; DNA: 21 BP.  
 AC ABR94128;  
 XX  
 DT 27-AUG-2002 (first entry)  
 DE  
 XX Endothelin-2 (EDN-2) SNP detection PCR primer #24.  
 DE  
 XX Endothelin: EDN; endothelin converting enzyme; ECE; endothelin receptor;  
 KW EDNR; signaling system; cardiovascular disease; coronary heart disease;  
 KW hypertension; atherosclerosis; angiogenesis; fatty acid metabolism;  
 KW diabetes; familial hypercholesterolemia; forensic marker;  
 KW transgenic animal; solid support; cardiovascular regulator; SNP;  
 KW single nucleotide polymorphism; PCR; primer; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200224747-A2.  
 XX  
 PD 28-MAR-2002.  
 XX  
 PF 31-AUG-2001; 2001MO-EP10087.  
 XX  
 PR 19-SEP-2000; 2000EP-0120123.  
 XX  
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX  
 PI Brinkmann U, Hoffmeyer S;  
 DR WPI: 2002-435060/46.  
 XX  
 XX Novel polynucleotide of the endothelin/endothelin converting  
 PT enzyme/receptors of endothelin and endothelin converting enzyme  
 PT signaling system associated with cardiovascular disease, useful for  
 PT treating the disease  
 XX  
 Claim 1; Page 57; 190pp; English.

The invention describes a polynucleotide (I) of the endothelin  
 CC (EDN)/endothelin converting enzyme (ECE)/receptors of EDN and ECE (EDNR)  
 CC signaling system which is associated with a cardiovascular disease. (I),  
 CC the gene encoding EDN, ECE or EDNR (II) or a vector (III) expressing (I)  
 CC or (II) is useful for producing cells capable of expressing a molecular  
 CC variant polypeptide which is associated with a cardiovascular disease.

CC (II), (III), the EDN, ECE or EDNR polypeptide, or a cell expressing  
CC a molecular variant gene comprising (I) is useful for identifying and  
CC obtaining a pro-drug or drug capable of modulating the activity of a  
CC molecular variant of a polypeptide of the EDN/EDNR/ECE signaling system  
CC or its gene product, or for identifying and obtaining an inhibitor of  
CC the activity of a molecular variant of a polypeptide of the EDN/EDNR/ECE  
CC signaling system or its gene product. The isolated proteins and  
CC polynucleotides encoding them are useful for preparation of a  
CC pharmaceutical composition for treating a cardiovascular disease such as  
CC coronary heart disease, hypertension, atherosclerosis, or related to  
CC abnormal angiogenesis or fatty acid metabolism e.g. diabetes and familial  
CC hypercholesterolaemia. The gene or a polynucleotide fragment of the  
CC EDN/ECE/EDNR signaling system are useful as forensic markers, for  
CC creating a transgenic animal and in creation of a solid support  
CC comprising polynucleotides, genes, vectors, polypeptides, antibodies or  
CC host cells of the invention. This sequence represents a PCR primer used  
CC to identify single nucleotide polymorphisms in DNA encoding  
CC cardiovascular regulator proteins of the EDN/ECE/EDNR signaling pathway.  
XX  
SQ Sequence 21 BP; 0 A; 10 C; 4 G; 7 T; 0 other;

Query Match 100.0%; Score 11; DB 24; Length 21;  
Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAGAGAG 11  
11 AGAGAGAGAG 1

RESULT 4  
AAT75373/c  
ID AAT75373 standard; cDNA; 22 BP.

AC AAT75373;

DT 24-DEC-1998 (first entry)

DE cDNA synthesis primer EGRI-6.

KW ss; human; RAD50; DNA repair; tumour suppression; cancer; Septin-2;

KM central nervous system; PCR; primer; amplification.

OS Synthetic.

PN WO9727284-A2.

PD 31-JUL-1997.

PF 24-JAN-1997; 97WO-US01299.

PR 17-JUL-1996; 96US-0687080.

PR 26-JAN-1996; 96US-0592126.

PA (GENE-) GENELABS TECHNOLOGIES INC.

PI Dolganov G;

DR WPI; 1997-393672/36.

PT Human tumour suppressor gene RAD50 - useful to detect  
PT predisposition to, decrease risk of and treat cancer, also Septin-2  
PT homologues

PS Example 1; Page 36; 155pp; English.

CC The primers AAT75354-T75378 were used to for cDNA synthesis in the  
CC method of the invention. Disclosed in the invention is human RAD50  
CC (RAD50) which is involved in DNA repair and has tumour suppression  
CC activity and can be used to detect predisposition to, decrease the risk  
CC of or treat cancers, e.g. acute myeloid leukaemia, myelodysplastic  
CC syndrome, therapy related myelodysplastic syndrome, therapy related  
CC acute myeloid leukaemia, refractory anaemia or refractory anaemia with

CC excess blasts. Also disclosed in this invention are human Septin-2  
CC homologues which may be used as targets for cancer therapies and central  
CC nervous system directed treatment methods, and to measure the  
CC proliferative potential of selected cell types.  
XX  
SQ Sequence 22 BP; 2 A; 12 C; 0 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 18; Length 22;  
Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAGAGAG 11  
15 AGAGAGAGAG 5

RESULT 5  
AAV59955/c  
ID AAV59955 standard; DNA; 22 BP.

AC AAV59955;

DT 25-NOV-1998 (first entry)

DE PCR primer EGRI-6 used to amplify EGRI-1 cDNA.

KW Human analogue; yeast RAD50; Drosophila Septin-2; Acyl-CoA synthetase;

KM immunomodulatory activity; Identification: activated T-cell; cytokine;

EGRI-1; PCR primer; ss.

OS Synthetic.

OS Homo sapiens.

PN WO9838306-A1.

PD 03-SEP-1998.

PF 27-FEB-1997; 97WO-US03159.

PR 27-FEB-1997; 97WO-US03159.

PA (GENE-) GENELABS TECHNOLOGIES INC.

PI Dolganov G;

DR WPI; 1998-481207/41.

PT Novel human immunomodulatory poly:peptide(s) - have homology to the  
PT yeast RAD50 or Drosophila Septin-2 proteins

PS Example 1; Page 27; 155pp; English.

CC PCR primers AAV59955-56 were used to identify cDNA encoding human  
CC cytokine EGRI-1 from different cDNA pools, to provide an estimate of  
CC the degree to which the cytokine transcript is present. mRNA was  
CC isolated from activated T-cells, and converted to cDNA prior to  
CC amplification. The specification describes sequences encoding human  
CC analogues of the yeast RAD50, the Drosophila Septin-2 and Acyl-CoA  
CC synthetase. The proteins have immunomodulatory activity. The nucleic  
CC acids and proteins can be used to identify activated T-cells in a sample  
CC population. They can also be used to isolate and identify sequences  
CC encoding other proteins or other compounds having immunomodulatory  
CC activity.

SQ Sequence 22 BP; 2 A; 12 C; 0 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 19; Length 22;  
Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAGAGAG 11  
15 AGAGAGAGAG 5

```

RESULT 6
AAQ39034
ID AAQ39034 standard; CDNA; 23 BP.
XX
AC AAQ39034;
XX
DT 22-JUL-1993 (first entry)
XX
DE Oligonucleotide identifier tag 3' primer.
XX
KW Oligonucleotide; identifier; tag; primer; oligomer; library; agonist;
KW identification; ligand; antagonist; cell membrane receptor; hormone;
KW toxin; venom; viral epitope; sugar; cofactor; peptide; drug; protein;
KW enzyme substrate; PCR; amplify; ss.
XX
OS Synthetic.
XX
PN WO9306121-A.
XX
PD 01-APR-1993.
XX
PF 16-SEP-1992; 92WO-US07815.
XX
PR 18-SEP-1991; 91US-0762522.
XX
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Barrett RW, Dower WJ, Gallop MA;
XX
DR WPI; 1993-117464/14.
XX
PT Synthetic oligomer library to facilitate identification of
PT oligomers with required properties - comprises several members,
PT each of which has an oligomer of monomer sequence linked to
PT identifier tag
XX
PS Disclosure; Page 37; 76pp; English.
XX
CC This sequence is a primer which was used in the preparation of the
CC sequences given in AAQ39030-33. The amplified oligonucleotides are
CC identifier tags. They were prepared using this primer and also a
CC primer corresponding to bases 7-28 of the sequence given in AAQ39030.
CC The tags were then sequenced using the primer sequences given in
CC AAQ39035-36. These identifier tags were used in the construction of
CC an oligomer library comprising >10,000,000 members. The use of the
CC identification tags on the oligomer members of the library
CC facilitates identification of oligomers with required properties.
CC Ligands which can be investigated include agonists and antagonists
CC for cell membrane receptors, toxins, venoms, viral epitopes, hormones,
CC sugars, cofactors, peptides, enzyme substrates, drugs and proteins.
XX
SQ Sequence 23 BP; 11 A; 0 C; 12 G; 0 U; 0 other;
Query Match 100.0%; Score 11; DB 14; Length 23;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGGAG 11
DB 10 AGAGAGAGGAG 20

```

```

XX
KW Cardiovascular disease; TagIB polymorphism; coronary heart disease; ss;
KW cholesteryl ester transfer protein; CERP; Tag I restriction site;
KW B1 allele; myocardial infarction; angina pectoris; coronary death;
KW coronary artery disease; coronary insufficiency; PCR primer; human.
XX
OS Homo sapiens.
XX
PN WO200185999-A1.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001; 2001WO-US15013.
XX
PR 11-MAY-2000; 2000US-203467P.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Ordovas JM, Schaefer EJ;
XX
DR WPI; 2002-082902/11.
XX
PF Assessing risk for the development of cardiovascular disease in an
PT individual, comprises analysing nucleic acid from the individual for
PT presence of TagIB polymorphism of cholesteryl ester transfer protein
PT gene
XX
PS Claim 13; Page 67; 72pp; English.
XX
CC The invention relates to a method for assessing risk for development of
CC cardiovascular disease, comprising analysing a nucleic acid from an
CC individual for the presence of TagIB polymorphism of cholesteryl ester
CC transfer protein (CERP) gene, and assessing the risk for development of
CC cardiovascular disease. The method involves determining whether the
CC individual is homozygous for, heterozygous for or does not possess the
CC TagIB polymorphism, whereby the absence of the polymorphism correlates
CC with the frequency and phenotypic expression of coronary heart disease.
CC The invention also includes PCR primers used for amplification of a
CC suitable section of the first intron of the CERP gene encompassing the
CC TagI restriction site of the B1 allele of the CERP gene, the presence of
CC the TagI restriction site being indicative of the absence of the TagIB
CC polymorphism. The cardiovascular diseases include myocardial infarction,
CC coronary artery disease, angina pectoris, coronary insufficiency and
CC coronary death. This sequence represents a PCR primer used in the method
CC of the invention.
XX
SQ Sequence 24 BP; 7 A; 7 C; 8 G; 2 T; 0 other;
Query Match 100.0%; Score 11; DB 24; Length 24;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGGAG 11
DB 10 AGAGAGAGGAG 20

```

```

RESULT 8
AAH37422/C
ID AAH37422 standard; DNA; 25 BP.
XX
AC AAH37422;
XX
DT 14-AUG-2001 (first entry)
XX
DE SNP specific lower PCR primer SEQ ID 218.
XX
KW Single nucleotide polymorphism; SNP; single nucleotide primer extension;
KW SNPE; genotyping; agammaglobulinaemia; diabetes insipidus; cancer;
KW Lesch-Nyhan syndrome; muscular dystrophy; familial hypercholesterolaemia;
KW polycystic kidney disease; osteogenesis imperfecta; autoimmune disease;
KW acute intermittent porphyria; rheumatoid arthritis; multiple sclerosis;
KW inflammation; forensic investigation; paternity analysis; PCR primer; ss.

```



XX OS Homo sapiens.  
XX PN WO200129262-A2.  
XX PD 26-APR-2001.  
XX PF 13-OCT-2000; 2000WO-US28436.  
XX PR 15-OCT-1999; 99US-0160096.  
XX PA (ORCH-) ORCHID BIOSCIENCES INC.  
XX PI Picoult-Newburg L, Pohl M;  
XX WPI; 2001-290930/30.  
XX DR  
XX PT New genotyping oligonucleotide, useful for detecting the presence,  
XX PT absence or identity of single polynucleotide polymorphism in a nucleic  
XX PT acid sample  
XX PS  
XX PS Claim 1; Page 51; 83pp: English.  
XX CC Sequences AAH37205 - AAH40944 represent PCR primers, single nucleotide  
XX CC primer extension (SNPE) primers, and the sequences of regions flanking  
XX CC sites of single nucleotide polymorphisms SNPs. The present invention  
XX CC includes kits for determining the presence or absence of a SNP, using the  
XX CC oligonucleotides of the invention. The PCR primers are used to amplify a  
XX CC SNP flanking sequence, the SNPE primer is used as a genotyping primer.  
XX CC The oligonucleotides are useful for genotyping a nucleic acid sample by  
XX CC performing a single-nucleotide primer extension reaction. The  
XX CC oligonucleotides are useful for determining the presence, absence or  
XX CC identity of a SNP and for genotyping nucleic acid samples, for e.g. to  
XX CC assess by association analysis the genotype of an individual or group of  
XX CC individuals, having a pathological phenotypic trait suspected of being  
XX CC caused by one or more SNPs. Phenotypic traits include diseases e.g.  
XX CC agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular  
XX CC dystrophy, familial hypercholesterolemia, polycystic kidney disease,  
XX CC osteogenesis imperfecta and acute intermittent porphyria. Phenotypic  
XX CC traits also include symptoms of or susceptibility to multifactorial  
XX CC disease of which a component is or may be genetic such as autoimmune  
XX CC diseases, including, rheumatoid arthritis, multiple sclerosis,  
XX CC inflammation, cancer, nervous system diseases and infection by pathogenic  
XX CC microorganism. The method is also useful in forensic investigations and  
XX CC paternity analysis. The present sequence represents a PCR primer specific  
XX CC for a human SNP containing DNA sequence.  
XX SQ  
XX SQ Sequence 25 BP; 4 A; 12 C; 0 G; 9 T; 0 other;  
XX  
XX Query Match 100.0%; Score 11; DB 22; Length 25;  
XX Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGGAG 11  
Db 22 AGAGAGAGGAG 12  
Db  
XX  
XX RESULT 9  
XX ID AAQ40991 standard; DNA; 31 BP.  
XX AAQ40991:  
XX  
XX AC AAQ40991:  
XX  
XX DT 17-DEC-2001 (updated)  
XX DT 06-OCT-1993 (first entry)  
XX  
XX DE Mono ADP-ribosyltransferase gene isolation sequence.  
XX  
XX KM Antisense technology; gene therapy; structure/function studies;  
XX KM production; preparation; prodn.; prepn.; antibody; NAD: arginine; ss.  
XX  
XX OS Synthetic.

XX XX USN7985698-N.  
XX PN 01-MAY-1993.  
XX PD 30-NOV-1992; 92US-0985698.  
XX PF 30-NOV-1992; 92US-0985698.  
XX PR 30-NOV-1992; 92US-0985698.  
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
XX PI Moss J, Nightingale MS, Okazaki I, Zolkiewska A;  
XX WPI; 1993-188869/23.  
XX DR  
XX XX Gene sequences encoding mono-ADP-ribosyl:transferase -  
XX PT specifically from rabbit skeletal muscle and human form, useful  
XX PT in antisense technology  
XX PS  
XX PS Example; Page 42; 46pp: English.  
XX  
XX CC The sequence is given in the specification but is not  
XX CC specifically defined.  
XX CC (Note: Revised entry submitted to correct the patent number format of  
XX CC US Government-owned NPLS applications to prevent clashes with ongoing US  
XX CC granted patent numbers. For further information please visit the Derwent  
XX CC web site at [www.derwent.com/dwpi/updates/npls-us.html](http://www.derwent.com/dwpi/updates/npls-us.html).)  
XX SQ  
XX SQ Sequence 31 BP; 12 A; 0 C; 19 G; 0 U; 0 other;  
XX  
XX Query Match 100.0%; Score 11; DB 14; Length 31;  
XX Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGGAG 11  
Db 5 AGAGAGAGGAG 15  
Db  
XX  
XX RESULT 10  
XX ID AAT91953/c  
XX AAT91953 standard; cDNA; 32 BP.  
XX  
XX AC AAT91953:  
XX  
XX DT 20-MAR-1998 (first entry)  
XX  
XX DE PCR primer 1 used to localize the antigenic peptide recognised by TIL586.  
XX  
XX KM Tyrosinase related protein 1 gene; TRP-1; gp75; tumour antigen;  
XX KM tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;  
XX KM alternative reading frame; cancer detection; pre-cancer detection;  
XX KM melanoma; PCR primer; ss.  
XX  
XX OS Synthetic.  
XX OS Homo sapiens.  
XX PN WO9729195-A2.  
XX PD 14-AUG-1997.  
XX PF 06-FEB-1997; 97WO-US02186.  
XX PR 04-OCT-1996; 96US-0725736.  
XX PR 09-FEB-1996; 96US-0595602.  
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX PI Rosenberg SA, Wang R;  
XX WPI; 1997-415349/38.  
XX  
XX KM Cancer antigen peptide(s) derived from the tyrosinase-related

PT protein 1 or 2 - useful for detecting, preventing or treating a  
 PT cancer in a mammal, especially melanoma  
 XX  
 PS Example 1; Page 34; 111pp; English.  
 XX  
 CC PCR primers AAT91953-54 were used to amplify the tyrosinase related  
 CC protein 1 gene (TRP-1 or gp75, AAT91944), from positions 1-667 in order  
 CC to localise the region containing the predominant T-cell epitope.  
 CC The PCR products contain an internal ATG codon in frame with the TRP-1  
 CC gene, located at 445 bp as a start codon for translation of a truncated  
 CC normal TRP-1 protein. However, neither of the PCR products conferred the  
 CC ability to stimulate cytokine secretion from TIL586, suggesting that the  
 CC epitope is located upstream of these fragments (see AAT91949-51). Novel  
 CC cancer peptides have been identified in TRP-1 and TRP-2. The nucleic  
 CC acids encoding the cancer peptides or TRP-2 can be used to detect a  
 CC cancer or pre-cancer in a mammal, especially by detecting the presence  
 CC of the alternative ORF 3 of the TRP-1 gene or the sequence encoding the  
 CC novel tumour antigen TRP-2. Vectors and recombinant viruses containing  
 CC antigen peptide encoding nucleic acids, antibodies raised against the  
 CC peptides, or the peptides themselves can be used to prevent or treat a  
 CC cancer in a mammal, especially a melanoma.  
 CC  
 SQ Sequence 32 BP; 7 A; 9 C; 7 G; 9 T; 0 other;  
 QY  
 Query Match 100.0%; Score 11; DB 18; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 AGAGAGAGGAG 11  
 |||||  
 29 AGAGAGAGGAG 19  
 Db  
 RESULT 11  
 AAF88029/c  
 ID AAF88029 standard; DNA: 36 BP.  
 XX  
 AC AAF88029;  
 XX  
 DT 16-JUL-2001 (first entry)  
 XX  
 DE Human Trp 1 PCR primer trp1-5/2-LIGA.  
 XX  
 KW Listeria: expression vector; tumor-associated antigen; Trp 1; Trp 2;  
 KW MelanA/MART-1; cytostatic; attenuated; immunotherapy; malignant melanoma;  
 KW pigmented tumor; malignant schwannoma; vaccination; tyrosinase;  
 KW antigen-presenting cell; PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200127295-A1.  
 XX  
 PD 19-APR-2001.  
 XX  
 PF 13-OCT-2000; 2000MO-DE03629.  
 XX  
 PR 14-OCT-1999; 99DE-1049594.  
 XX  
 PA (DEKR-) DEUT KREBSFORSCHUNGSENTRUM.  
 XX  
 PI Schandendorf D, Paschen A, Chakraborty T, Dommann E;  
 DR WPI: 2001-282041/29.  
 XX  
 PT Listeria expression vector for immunotherapy, particularly of malignant  
 PT melanoma, comprises a DNA sequence encoding tumor-associated antigens -  
 PS Examples: Page 16; 41pp; German.  
 XX  
 CC This invention describes a novel Listeria expression vector (A) for  
 CC immunotherapy which comprises a promoter (P) functional in Listeria,  
 CC operably linked to a DNA sequence (I) encoding one of the  
 CC tumor-associated antigens (II) human tyrosinase, Trp 1 or 2, or

CC MelanA/MART-1. The products of the invention have cytostatic activity.  
 CC Recombinant attenuated Listeria containing (A) are useful for  
 CC immunotherapy (prophylactic, adjuvant or therapeutic), specifically of  
 CC malignant melanoma (but also other pigmented tumors such as malignant  
 CC schwannoma), particularly as a replacement for radiotherapy. Using  
 CC attenuated Listeria as carrier for (A) provides a simple way of  
 CC vaccination, since antigen-presenting cells acquire tumor-associated  
 CC antigens by natural infection, eliminating the need for labor-intensive  
 CC ex vivo modification of autologous cells. This sequence represents the  
 CC PCR primer trp1-5/2-LIGA used in the amplification of the human  
 CC Trp 1 which is described in the method of the invention.  
 XX  
 SQ Sequence 36 BP; 8 A; 13 C; 5 G; 10 T; 0 other;  
 QY  
 Query Match 100.0%; Score 11; DB 22; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 AGAGAGAGGAG 11  
 |||||  
 35 AGAGAGAGGAG 25  
 Db  
 RESULT 12  
 AAZ95968  
 ID AAZ95968 standard; DNA: 40 BP.  
 XX  
 AC AAZ95968;  
 XX  
 DT 10-APR-2000 (first entry)  
 XX  
 DE Polynucleotide sequence including binding site for BamHI.  
 XX  
 KW Ligand binding; restriction enzyme; nucleic acid determination;  
 KW pharmaceutical; BamHI. ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO963077-A2.  
 XX  
 PD 09-DEC-1999.  
 XX  
 PF 04-JUN-1999; 99MO-US12516.  
 XX  
 PR 04-JUN-1998; 98US-0087905.  
 XX  
 PR 03-JUN-1999; 99US-0324672.  
 XX  
 PA (TMTE-) TM TECHNOLOGIES INC.  
 XX  
 PI Lane MJ, Benight AS, Faldasz BD;  
 DR WPI: 2000-116369/10.  
 XX  
 PT Modulating polynucleotide ligand binding site affinity using  
 PT determination of the flanking duplex sequences -  
 PS Example 1; Page 40; 62pp; English.  
 XX  
 CC The invention provides a method for determining the sequence of  
 CC polynucleotide flanking regions that modulate ligand binding  
 CC characteristics of an adjacent binding site. The method comprises:  
 CC (i) providing a number of different duplex polynucleotides, each having  
 CC the same polynucleotide ligand binding site and a randomly synthesised  
 CC sequence flanking the binding site; (ii) exposing the duplex to a ligand  
 CC selective for the binding site; (iii) isolating duplexes which bind or  
 CC do not bind the ligand, and (iv) determining the nucleotide composition  
 CC of the flanking duplex sequence by sequencing the duplex sequence  
 CC adjacent to the binding site. The invention is used to modulate the  
 CC ligand-binding characteristics of any nucleotide sequence. The invention  
 CC is less costly and more efficient than prior art techniques that moderate  
 CC ligand binding using small molecule pharmaceuticals. Sequences  
 CC AAZ95762-256170 represent polynucleotide sequences including the binding  
 CC site for the restriction enzyme BamHI and used in the course of the

CC Invention.  
XX  
SQ Sequence 40 BP; 12 A; 7 C; 14 G; 7 T; 0 other;  
Query Match 100.0%; Score 11; DB 21; Length 40;  
Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGAGAG 11  
DB 28 AGAGAGAGAGAG 38  
RESULT 13  
AAx81695  
ID AAx81695 standard; DNA; 43 BP.  
AAx81695;  
AC  
XX  
XX  
DT 03-SEP-1999 (first entry)  
DE Primer used in a novel method of DNA sequencing.  
XX  
XX  
XX Extension region; terminal nucleotide; template arranging segment;  
KW complexity decreasing nucleotide; DNA sequencing; primer; ss.  
XX  
OS Synthetic.  
XX JP1151092-A.  
XX  
PD 08-JUN-1999.  
XX  
PD 24-AUG-1998; 98JP-0237840.  
PF  
PR 22-AUG-1997; 97US-0916120.  
XX  
XX (LYNX-) LYNX THERAPEUTICS INC.  
PA WPI: 1999-388481/33.  
XX  
XX Extension of DNA using rolling primer - useful as DNA sequencing  
PT method  
XX  
PS Disclosure; Page 26; 32pp; Japanese.  
XX  
XX The specification describes a method for the determination of the  
CC nucleotide sequence of a polynucleotide. The method comprises providing  
CC a set of primers in which each primer has an extension region containing  
CC a terminal nucleotide, a template arranging segment and at least one  
CC complexity-decreasing nucleotide, forming a template containing  
CC primer-combining sites and the polynucleotide in which the  
CC primer-combining sites are complementary to at least one primer of the  
CC set, forming an amplicon from the template by amplifying a  
CC double-stranded DNA formed selectively by extending the primer from  
CC the set in which the extending region forms a double-strand completely  
CC matched to primer-combining sites of the template, identifying the  
CC terminal nucleotide of the extending region of the primer by an  
CC identification of the amplicon, shifting the primer-combining sites by  
CC one nucleotide to the direction of extension by varying the  
CC primer-combining sites of the template, and repeating this until the  
CC nucleotide sequence of the polynucleotide is determined. The method can  
CC be used for DNA sequencing. The present sequence represents a primer  
CC used to demonstrate the invention.  
XX  
SQ Sequence 43 BP; 15 A; 2 C; 17 G; 1 T; 8 other;  
Query Match 100.0%; Score 11; DB 20; Length 43;  
Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGAGAG 11  
DB 13 AGAGAGAGAGAG 23

RESULT 14  
AAx77247  
ID AAx77247 standard; cDNA; 51 BP.  
XX  
XX  
AC AAx77247;  
XX  
XX  
DT 16-NOV-2000 (first entry)  
DE Human clone c943979039 polymorphic site, SEQ ID NO:930.  
XX  
XX  
KW Human; single nucleotide polymorphism; SNP; chromosome 2;  
KW detection; identification; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH variation replace (26,A)  
FT /\*tag=a  
XX  
XX WO200029623-A2.  
XX  
XX  
PD 25-MAY-2000.  
XX  
XX 17-NOV-1999; 99WO-US27293.  
PF  
PR 17-NOV-1998; 98US-0109024.  
PR 16-NOV-1999; 99US-0109024.  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX Shinketsu RA, Leach MD;  
PI  
XX WPI: 2000-387826/33.  
XX  
XX Human nucleic acids containing single nucleotide polymorphisms, useful  
PT for treating a subject suffering, or at risk from a pathology due to  
PT the presence of a sequence polymorphism -  
XX  
XX Claim 1; Page 438; 543pp; English.  
XX  
XX Sequences AAx76318-A77509 represent 1192 human nucleic acid sequences  
CC which contain single nucleotide polymorphisms (SNPs). Sequences 1 to  
CC 1112 (AAx76318-A77429) are consecutive pairs of nucleotides which  
CC contain silent SNPs. Sequences 1113 to 1192 (AAx77430-A77509) are  
CC consecutive pairs of nucleotides containing SNPs which result in changes  
CC in the corresponding amino acid sequences (AAx11749-B11828). The SNPs in  
CC sequences 1113 to 1128 (AAx77430-A77445) lead to conservative amino acid  
CC changes, while those in sequences 1129 to 1186 (AAx77446-A77503) result  
CC in non-conservative changes. The SNPs in sequences 1187 to 1192  
CC (AAx77504-A77509) generate frameshift mutations. The invention also  
CC relates to a method of detecting a polymorphic site in a nucleic acid and  
CC a method of determining the relatedness of two nucleic acids. It also  
CC encompasses peptides containing polymorphic sites, antibodies raised  
CC against such peptides, and a method of detecting polymorphic  
CC proteins/peptides using the antibodies. The nucleic acids are useful for  
CC gene therapy of an individual having, suspected of having, or at risk of  
CC developing a pathological condition due to the presence of a sequence  
CC polymorphism. Such treatment would comprise administration of the  
CC wild-type nucleic acid sequence. Antibodies raised against polymorphic  
CC peptides can also be used in the treatment of such individuals.  
XX  
SQ Sequence 51 BP; 13 A; 20 C; 12 G; 6 T; 0 other;  
Query Match 100.0%; Score 11; DB 21; Length 51;  
Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGAGAG 11  
DB 16 AGAGAGAGAGAG 26

Search completed: July 28, 2003, 11:28:03  
 Job time : 6.2668 secs

```

RESULT 15
ID ABN40086/C
AC ABN40086 standard; DNA; 60 BP.
XX
AC ABN40086;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:12834.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KM splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN W0200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI WPI; 2002-257383/30.
XX
DR
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX
PS Example 1; SEQ ID 12834; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 60 BP; 5 A; 25 C; 8 G; 22 T; 0 other;

```

Query Match 100.0%; Score 11; DB 24; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AGAGAGAGAGAG 11
   |||||||
Db 45 AGAGAGAGAG 35

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 10:00:14 ; Search time 0.694102 Seconds  
(without alignments)  
4860.160 Million cell updates/sec

Title: US-09-902-772-1\_COPY\_650\_660

Sequence: 1 agagagagagag 11

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	11	100.0	22	1	US-08-592-126-24
C 2	11	100.0	22	2	US-08-687-080-24
C 3	11	100.0	23	1	US-07-946-239-5
C 4	11	100.0	23	3	US-09-151-467-5
C 5	11	100.0	23	3	US-09-036-599-10
C 6	11	100.0	23	4	US-09-256-838-5
C 7	11	100.0	23	5	PCR-US92-07815-5
C 8	11	100.0	30	4	US-09-147-915-6
C 9	11	100.0	32	2	US-08-599-602-42
C 10	11	100.0	32	3	US-09-197-816-42
C 11	11	100.0	36	2	US-08-863-639A-26
C 12	11	100.0	43	1	US-08-560-313A-8
C 13	11	100.0	43	1	US-08-611-155B-12
C 14	11	100.0	43	2	US-08-916-120A-14
C 15	11	100.0	63	4	US-08-951-822-12
C 16	11	100.0	63	4	US-09-173-043-26
C 17	11	100.0	63	4	US-09-368-951-12
C 18	11	100.0	90	1	US-08-314-536-34
C 19	11	100.0	90	1	US-08-320-982-34
C 20	11	100.0	90	3	US-08-819-037-34
C 21	11	100.0	90	4	US-09-045-940-34
C 22	11	100.0	104	1	US-08-314-536-34
C 23	11	100.0	104	1	US-08-320-982-31
C 24	11	100.0	104	3	US-08-819-037-31
C 25	11	100.0	104	4	US-09-045-940-31
C 26	11	100.0	111	1	US-08-488-470A-9
C 27	11	100.0	111	1	US-08-488-470A-10

C 28	11	100.0	111	1	US-07-946-239-1	Sequence 1, Appl
C 29	11	100.0	111	1	US-07-946-239-2	Sequence 2, Appl
C 30	11	100.0	111	1	US-08-484-505A-9	Sequence 9, Appl
C 31	11	100.0	111	1	US-08-484-505A-10	Sequence 10, Appl
C 32	11	100.0	111	3	US-09-151-467-1	Sequence 1, Appl
C 33	11	100.0	111	3	US-09-151-467-2	Sequence 2, Appl
C 34	11	100.0	111	3	US-09-036-599-5	Sequence 5, Appl
C 35	11	100.0	111	3	US-09-036-599-6	Sequence 6, Appl
C 36	11	100.0	111	4	US-09-078-403A-9	Sequence 9, Appl
C 37	11	100.0	111	4	US-09-078-403A-10	Sequence 10, Appl
C 38	11	100.0	111	4	US-09-256-838-1	Sequence 1, Appl
C 39	11	100.0	111	4	US-09-256-838-2	Sequence 2, Appl
C 40	11	100.0	111	5	PCR-US92-07815-1	Sequence 1, Appl
C 41	11	100.0	111	5	PCR-US92-07815-2	Sequence 2, Appl
C 42	11	100.0	115	1	US-08-488-470A-11	Sequence 11, Appl
C 43	11	100.0	115	1	US-08-488-470A-12	Sequence 12, Appl
C 44	11	100.0	115	1	US-07-946-239-3	Sequence 3, Appl
C 45	11	100.0	115	1	US-07-946-239-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-592-126-24/c  
Sequence 24, Application US/08592126  
Patent No. 5821091  
GENERAL INFORMATION:  
APPLICANT: Gregory Dolganov  
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory  
TITLE OF INVENTION: Polypeptides  
NUMBER OF SEQUENCES: 151  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Denlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,126  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO.: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
INDIVIDUAL SOURCE:  
US-08-592-126-24  
INDIVIDUAL ISOLATE: Primer EGR1-6  
Query Match 100.0%; Score 11; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11

Db 15 AGAGAGAGAG 5

## RESULT 2

US-08-687-080-24/C  
Sequence 24, Application US/08687080  
Patent No. 5963427  
GENERAL INFORMATION:  
APPLICANT: Gregory Dolganov  
TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof  
NUMBER OF SEQUENCES: 175  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dellinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,080  
FILING DATE: 17-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/592,126  
FILING DATE: 26-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0111.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Primer EGR11-6  
US-08-687-080-24

Query Match 100.0%; Score 11; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
Db 15 AGAGAGAGAG 5

RESULT 3  
US-07-946-239-5  
Sequence 5, Application US/07946239  
Patent No. 5770358  
GENERAL INFORMATION:  
APPLICANT: DOWER, WILLIAM J  
APPLICANT: BARRETT, RONALD W  
APPLICANT: GALLOP, MARK A  
APPLICANT: NEEDLES, MICHAEL C  
TITLE OF INVENTION: METHOD OF SYNTHESIZING DIVERSE  
NUMBER OF SEQUENCES: 16  
COLLECTIONS OF OLIGOMERS  
CORRESPONDENCE ADDRESS:

ADDRESSEE: TOWNSEND AND TOWNSEND  
STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94105

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/946,239  
FILING DATE: 19920916  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11509-36-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-5043  
TELEFAX: 415-543-9600

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA

US-07-946-239-5

Query Match 100.0%; Score 11; DB 1; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
Db 10 AGAGAGAGAG 20

RESULT 4  
US-09-151-467-5  
Sequence 5, Application US/09151467  
Patent No. 6140493  
GENERAL INFORMATION:  
APPLICANT: DOWER, WILLIAM J  
APPLICANT: BARRETT, RONALD W  
APPLICANT: GALLOP, MARK A  
APPLICANT: NEEDLES, MICHAEL C  
TITLE OF INVENTION: METHOD OF SYNTHESIZING DIVERSE  
NUMBER OF SEQUENCES: 16  
COLLECTIONS OF OLIGOMERS  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND AND TOWNSEND  
STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/151,467  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/946,239  
FILING DATE: 1992-09-16  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.

REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11509-36-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-151-467-5

Query Match 100.0%; Score 11; DB 3; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGAGAGAGGAG 11  
|||||  
10 AGAGAGAGGAG 20

Db

RESULT 5  
US-09-036-599-10  
Sequence 10, Application US/09036599  
Patent No. 6143497  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Gallop, Mark A.  
TITLE OF INVENTION: Method of Synthesizing Diverse  
TITLE OF INVENTION: Collections of Oligomers  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESS: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/036,599  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/762,522  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11509-36  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-5043  
TELEFAX: 415-326-2400  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-036-599-10

Query Match 100.0%; Score 11; DB 3; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGAGAGAGGAG 11  
|||||  
10 AGAGAGAGGAG 20

Db

RESULT 6  
US-09-256-838-5  
Sequence 5, Application US/09256838  
Patent No. 6416949  
GENERAL INFORMATION:  
APPLICANT: DOWER, WILLIAM J  
BARRETT, RONALD W  
GALLOP, MARK A  
NEDELS, MICHAEL C  
TITLE OF INVENTION: METHOD OF SYNTHESIZING DIVERSE  
TITLE OF INVENTION: COLLECTIONS OF OLIGOMERS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESS: TOWNSEND AND TOWNSEND  
STREET: 1 MARKET PLAZA, STEWART TOWER, SUITE 2000  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/256,838  
FILING DATE: 24-Feb-1999  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/946,239  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11509-36-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-256-838-5

Query Match 100.0%; Score 11; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGAGAGAGGAG 11  
|||||  
10 AGAGAGAGGAG 20

Db

RESULT 7  
PCT-US92-07815-5  
Sequence 5, Application PC/TUS9207815  
GENERAL INFORMATION:  
APPLICANT: DOWER, WILLIAM J  
BARRETT, RONALD W  
APPLICANT: GALLOP, MARK A  
APPLICANT: NEDELS, MICHAEL C  
TITLE OF INVENTION: METHOD OF SYNTHESIZING DIVERSE  
TITLE OF INVENTION: COLLECTIONS OF OLIGOMERS  
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND AND TOWNSEND  
STREET: 1 MARKET PLAZA, STEWART TOWER, SUITE 2000  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/07815  
FILING DATE: 19920916  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11509-36-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
PCT-US92-07815-5

Query Match 100.0%; Score 11; DB 5; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
|||||  
DB 10 AGAGAGAGGAG 20

RESULT 8  
US-09-147-915-6/c  
Sequence 6, Application US/09147915A  
Patent No. 6184034  
GENERAL INFORMATION:  
APPLICANT: Eastman, Alan  
APPLICANT: Krieser, Ronald  
TITLE OF INVENTION: Deoxyribonuclease II proteins and cDNAs  
FILE REFERENCE: DC-0097  
CURRENT APPLICATION NUMBER: US/09/147,915A  
CURRENT FILING DATE: 1999-03-23  
EARLIER APPLICATION NUMBER: PCT/US97/18262  
EARLIER FILING DATE: 1997-10-09  
EARLIER APPLICATION NUMBER: 60/028,539  
EARLIER FILING DATE: 1996-10-15  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 30  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURe:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-147-915-6

Query Match 100.0%; Score 11; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
|||||  
DB 29 AGAGAGAGGAG 19

RESULT 9  
US-08-599-602-42/c  
Sequence 42, Application US/08599602  
Patent No. 5840839  
GENERAL INFORMATION:  
APPLICANT: WANG, R-F.; ROSENBERG, S.A.  
TITLE OF INVENTION: ALTERNATIVE OPEN READING FRAME DNA OF A  
TITLE OF INVENTION: NORMAL GENE AND A NOVEL HUMAN CANCER ANTIGEN ENCODED  
TITLE OF INVENTION: THEREIN  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5-INCH, 1.44MB STORAGE  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/599,602  
FILING DATE: 09 FEB 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KATHRYN M. BROWN  
REGISTRATION NUMBER: 34,556  
REFERENCE/DOCKET NUMBER: 2026-4209  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
US-08-599-602-42

Query Match 100.0%; Score 11; DB 2; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
|||||  
DB 29 AGAGAGAGGAG 19

RESULT 10  
US-09-197-816-42/c  
Sequence 42, Application US/09197816  
Patent No. 6087110  
GENERAL INFORMATION:  
APPLICANT: WANG, R-F.; ROSENBERG, S.A.  
TITLE OF INVENTION: ALTERNATIVE OPEN READING FRAME DNA OF A  
TITLE OF INVENTION: NORMAL GENE AND A NOVEL HUMAN CANCER ANTIGEN ENCODED  
TITLE OF INVENTION: THEREIN  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154



COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5-INCH, 1.44MB STORAGE  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/197,816  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/599,602  
FILING DATE: 09 FEB 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: KATHRYN M. BROWN  
REGISTRATION NUMBER: 34,556  
REFERENCE/DOCKET NUMBER: 2026-4209  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
US-09-197-816-42

Query Match 100.0%; Score 11; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
DB 29 AGAGAGAGAG 19

RESULT 11  
US-08-863-639A-26  
Sequence 26, Application US/08863639A  
Patent No. 5981185  
GENERAL INFORMATION:  
APPLICANT: Matson, Robert S.  
APPLICANT: Coassin, Peter J.  
APPLICANT: Rampal, Jang B.  
APPLICANT: Caskey, C. T.  
TITLE OF INVENTION: OLIGONUCLEOTIDE REPEAT ARRAYS  
NUMBER OF SEQUENCES: 95  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, 9th Floor  
CITY: Pasadena  
STATE: CA  
COUNTRY: USA  
ZIP: 91101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Corel wordperfect 8 version  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/863,639A  
FILING DATE: May 28, 1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Joseph E. Mueh  
REGISTRATION NUMBER: 20,532  
REFERENCE/DOCKET NUMBER: 11859-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (626) 796-4000

TELEFAX: (626) 795-6321  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Other nucleic acid  
US-08-863-639A-26

Query Match 100.0%; Score 11; DB 2; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
DB 18 AGAGAGAGAG 28

RESULT 12  
US-08-560-313A-8  
Sequence 8, Application US/08560313A  
Patent No. 5763175  
GENERAL INFORMATION:  
APPLICANT: Sydney Brenner  
TITLE OF INVENTION: Simultaneous Sequencing of Tagged Polynucleotides  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Stephen C. Macevicz, Lynx Therapeutics, Inc.  
STREET: 3832 Bay Center Place  
CITY: Hayward  
STATE: California  
COUNTRY: USA  
ZIP: 94545  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: Power Macintosh  
OPERATING SYSTEM: Macintosh OS ver. 7.5.2  
SOFTWARE: Microsoft Word, vers. 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/560,313A  
FILING DATE: 17-NOV-95  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Stephen C. Macevicz  
REGISTRATION NUMBER: 30,285  
REFERENCE/DOCKET NUMBER: sstllus  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 670-9365  
TELEFAX: (510) 670-9302  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-560-313A-8

Query Match 100.0%; Score 11; DB 1; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
DB 13 AGAGAGAGAG 23

RESULT 13  
US-08-611-155B-12  
Sequence 12, Application US/0861155B

Patent No. 5780231  
GENERAL INFORMATION:  
APPLICANT: Sydney Brenner  
TITLE OF INVENTION: DNA Extension and Analysis with Rolling Primers  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Stephen C. Macevitz, Lynx Therapeutics, Inc.  
STREET: 3832 Bay Center Place  
CITY: Hayward  
STATE: California  
COUNTRY: USA  
ZIP: 94545  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: Windows 3.1  
SOFTWARE: Microsoft Word, vers. 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/611,155B  
FILING DATE: 05-MAR-96  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/560,313  
FILING DATE: 17-NOV-95  
ATTORNEY/AGENT INFORMATION:  
NAME: Stephen C. Macevitz  
REGISTRATION NUMBER: 30,285  
REFERENCE/DOCKET NUMBER: strplus  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 670-9365  
TELEFAX: (510) 670-9302  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-611-155B-12  
Query Match 100.0%; Score 11; DB 1; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGAG 11  
DB 13 AGAGAGAGAG 23  
RESULT 14  
US-08-916-120A-14  
Sequence 14, Application US/08916120A  
Patent No. 5962228  
GENERAL INFORMATION:  
APPLICANT: Sydney Brenner  
TITLE OF INVENTION: DNA Extension and Analysis with Rolling Primers  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Stephen C. Macevitz, Lynx Therapeutics, Inc.  
STREET: 3832 Bay Center Place  
CITY: Hayward  
STATE: California  
COUNTRY: USA  
ZIP: 94545  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Microsoft Word, vers. 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/916,120A  
FILING DATE: 22-AUG-97  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/611,155  
FILING DATE: 05-MAR-96  
ATTORNEY/AGENT INFORMATION:  
NAME: Stephen C. Macevitz  
REGISTRATION NUMBER: 30,285  
REFERENCE/DOCKET NUMBER: 811-01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 670-9365  
TELEFAX: (510) 670-9302  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-916-120A-14  
Query Match 100.0%; Score 11; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGAG 11  
DB 13 AGAGAGAGAG 23  
RESULT 15  
US-08-951-822-12  
Sequence 12, Application US/08951822A  
Patent No. 5989866  
GENERAL INFORMATION:  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Conklin, Darrell C.  
APPLICANT: Raymond, Fenella  
APPLICANT: Bukowski, Thomas R.  
APPLICANT: Holderman, Susan D.  
APPLICANT: Hansen, Birgit  
APPLICANT: Sheppard, Paul O.  
TITLE OF INVENTION: NOVEL FGF HOMOLOGS  
FILE REFERENCE: 96-20  
CURRENT APPLICATION NUMBER: US/08/951,822A  
CURRENT FILING DATE: 1997-10-16  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 63  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: oligonucleotide primer ZC15131  
US-08-951-822-12  
Query Match 100.0%; Score 11; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGAG 11  
DB 17 AGAGAGAGAG 27  
Search completed: July 28, 2003, 12:16:32  
Job time: 1.6941 secs

Seq primer: CACACAGGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 54.  
 Location/Qualifiers  
 1..54  
 source

/organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="U082M0282F14"  
 /clone\_1lb="Mouse 10kb plasmid U082M library"  
 /sex="Female"  
 /lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 1 a 21 c 5 g 27 t  
 ORIGIN

Query Match 100.0%; Score 11; DB 17; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+04;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGAGAGAGAG 11  
 |||||  
 Db 15 AGAGAGAGAG 5

Search completed: July 28, 2003, 12:14:56  
 Job time : 29.0947 secs

the site of insertion. Details of the protocols used can be found at [http://signal.salk.edu/cdna\\_protocols.html](http://signal.salk.edu/cdna_protocols.html)

BASE COUNT

4 a 13 c 12 g 17 t

ORIGIN

Query Match 100.0%; Score 11; DB 17; Length 46;  
Best Local Similarity 100.0%; Pred. No. 5.1e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
|||||

DB 28 AGAGAGAGAG 18

RESULT 13

LOCUS

BF651086 52 bp mRNA linear EST 20-DEC-2000

DEFINITION

clone NF100H09EC1078 Elicited cell culture Medicago truncatula cDNA

ACCESSION

BF651086

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)  
Contact: Dixon RA  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7302  
Fax: 580 221 7380  
Email: radixon@noble.org  
Insert Length: 52 Std Error: 0.00  
Plate: 100 row: H column: 09  
Seq primer: TCACACAGAGAAACAGCTATGAC.  
Location/Qualifiers

FEATURES

SOURCE

1. 52  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone="NF100H09EC"  
/clone\_id="Elicited cell culture"  
/tissue-type="Cell cultures derived from root tissues"  
/dev-stage="Cell suspensions were subcultured every 14 days. Cells were induced six days after subculture."  
/note="Vector: Lambda Zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."  
time point were pooled and used for mRNA isolation."  
BASE COUNT 4 a 21 c 2 g 25 t  
ORIGIN

Query Match 100.0%; Score 11; DB 12; Length 52;  
Best Local Similarity 100.0%; Pred. No. 5.3e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
|||||

DB 43 AGAGAGAGAG 33

RESULT 14  
BM395183/c 54 bp mRNA linear EST 17-JAN-2002

DEFINITION 50072-2-7-H10.f.2 Chilcoat/Turkewitz cDNA (large fraction)

Tetrahymena thermophila cDNA, mRNA sequence.

ACCESSION

BM395183

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2002)  
Contact: Turkewitz AP  
Molecular Genetics and Cell Biology  
University of Chicago  
920 E. 58th Street, Chicago, IL 60637, USA  
Tel: 773 702 4374  
Fax: 773 702 3172  
Email: apturkew@midway.uchicago.edu  
Seq primer: T3.  
Location/Qualifiers

1. 54  
/organism="Tetrahymena thermophila"  
/strain="CU428.1"  
/db\_xref="taxon:5911"  
/clone\_id="Chilcoat/Turkewitz cDNA (large fraction)"  
/note="Vector: Bluescript SK+; Details on library preparation can be found in Chilcoat and Turkewitz (2001) Proc. Natl. Acad. Sci USA, 98: 8709-8713."  
BASE COUNT 0 a 34 c 0 g 20 t  
ORIGIN

Query Match 100.0%; Score 11; DB 13; Length 54;  
Best Local Similarity 100.0%; Pred. No. 5.4e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
|||||

DB 44 AGAGAGAGAG 34

RESULT 15  
LOCUS A2996539 54 bp DNA linear GSS 27-APR-2001

DEFINITION ZM0282F14R Mouse 10kb plasmid U06C2M library Mus musculus genomic

ACCESSION A2996539

VERSION A2996539

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0282 row: F column: 14

**KEYWORDS** EST.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**REFERENCE** 1 (bases 1 to 40)  
**AUTHORS** NCI/NIH-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**TITLE** National Cancer Institute / National Institute of Dental Research, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
**JOURNAL** Unpublished (1997)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: [cgapds-remail.nih.gov](mailto:cgapds-remail.nih.gov)  
 Tissue Procurement: John Ensley, M.D.  
 CDNA Library Preparation: Stratagene, Inc.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: [www-bio.1nl.gov/bdpr/image/image.html](http://www-bio.1nl.gov/bdpr/image/image.html)

**FEATURES**  
 source  
 1. 40  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1lb="NCI-CGAP\_HN3"  
 /tissue\_type="squamous cell carcinoma from base of tongue"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /note="Organ: tongue; Vector: Bluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; cloned unidirectionally. Primer: 5' GAATTCGCGCAGAG 3' adaptor sequence: 5' (GA )10ACTAGTCGAGGTTTTTTTTTTTTTTT 3' "

**BASE COUNT**  
 ORIGIN  
 16 a 3 c 15 g 6 t

**Query Match** 100.0%; Score 11; DB 9; Length 40;  
**Best Local Similarity** 100.0%; Pred. No. 4.9e+04;  
**Matches** 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 1 AGAGAGAGAG 11  
 |||||  
**Db** 15 AGAGAGAGAG 25

**RESULT 11**  
**LOCUS** B1252040 40 bp mRNA linear EST 17-JUN-2001  
**DEFINITION** 602952695F1 NIH\_MGC\_100 Homo sapiens CDNA clone IMAGE:5087039 5', mRNA sequence.  
**ACCESSION** B1252040  
**VERSION** B1252040.1 GI:14802111  
**KEYWORDS** EST.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**REFERENCE** 1 (bases 1 to 40)  
**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: [cgapds-remail.nih.gov](mailto:cgapds-remail.nih.gov)  
 Tissue Procurement: CGAP (Stanford)  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

<http://image.1nl.gov>  
 Plate: LICM1843 row: m column: 24  
 High quality sequence stop: 19.

**FEATURES**  
 source  
 1. 40  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1lb="IMAGE:5087039"  
 /tissue\_type="NIH MGC 100"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pOTB; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

**BASE COUNT**  
 ORIGIN  
 18 a 1 c 20 g 1 t

**Query Match** 100.0%; Score 11; DB 13; Length 40;  
**Best Local Similarity** 100.0%; Pred. No. 4.9e+04;  
**Matches** 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 1 AGAGAGAGAG 11  
 |||||  
**Db** 13 AGAGAGAGAG 23

**RESULT 12**  
**LOCUS** BH812871/c 46 bp DNA linear GSS 02-MAY-2002  
**DEFINITION** SALK\_063356 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK\_063356, DNA sequence.  
**ACCESSION** BH812871  
**VERSION** BH812871.1 GI:20391326  
**KEYWORDS** GSS.  
**SOURCE** thale cress.  
**ORGANISM** Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 46)  
 Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.  
 A Sequence-indexed Library of Insertion Mutations in the Arabidopsis Genome  
 Unpublished (2001)  
**COMMENT** Contact: Joseph R. Ecker  
 Salk Institute Genomic Analysis Laboratory (SIGNAL)  
 The Salk Institute for Biological Studies  
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 Tel: 858 453 4100 x1752  
 Fax: 858 558 6379  
 Email: [ecker@salk.edu](mailto:ecker@salk.edu)  
 This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At5g40430.  
 Class: TDNA tagged.

**FEATURES**  
 source  
 1. 46  
 /organism="Arabidopsis thaliana"  
 /strain="Columbia 0"  
 /db\_xref="taxon:3702"  
 /clone\_1lb="SALK\_063356"  
 /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at

AA922878  
; mRNA sequence.

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (911473114[gb]AF12072.1), a copy number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT	2 a	19 c	0 g	12 t
ORIGIN	.			

Query Match	100.0%;	Score 11;	DB 17;	Length 33;
Best Local Similarity	100.0%;	Pred. No. 4.6e+04;		
Matches 11;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	1	AGAGAGAGGAG	11
Db	23	AGAGAGAGGAG	13

RESULT 7	LOCUS	DEFINITION
A2351541/c	36 bp	DNA linear GSS 29-SEP-2000
A2351541	1M0089G11R Mouse 10kb plasmid	U06C1M library Mus musculus genomic
	clone U06C1M0089G11 R,	DNA sequence.

ALPHABET	035.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE AUTHORS	TITLE	JOURNAL COMMENT
1 (Pages 1 to 36) Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausen, A. and Wright, D., Weiss, R.	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts	Unpublished (2000) Contact: Robert B. Weiss

PM. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
 84112, USA  
 Tel.: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0089 row: 6 column: 11  
 Seq primer: CACACAGAAACACACTGTGACC  
 Class: plasmid ends  
 High quality sequence stop: 36.  
 Location/Qualifiers  
 1..36

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0089G11"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42NV; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA

```

TDNA.  
Class: TDNA tagged.  
Location/Qualifiers  
1. 24  
/organism="Arabidopsis thaliana"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="SALK\_019058.23.05.x"

/note="PCR was performed on Arabidopsis thaliana TDNA insertion lines" each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

BASE COUNT  
ORIGIN  
2 a 14 c 0 g 8 t

Query Match 100.0%; Score 11; DB 17; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.1e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
Db 14 AGAGAGAGGAG 4

#### RESULT 4

LOCUS A2635993 25 bp DNA linear GSS 13-DEC-2000  
DEFINITION IM0493E20R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0493E20 R, DNA sequence.

ACCESSION A2635993  
VERSION A2635993.1 GI:11758183

KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 25)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunne@genetics.utah.edu  
Insert length: 10000 Std Error: 0.00  
Plate: 0493 row: E column: 20  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 25.

FEATURES  
source 1. 25  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0493E20"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (g147321149b1AF129072.1) a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
11 a 0 c 14 g 0 t

Query Match 100.0%; Score 11; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 4.2e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
Db 5 AGAGAGAGGAG 15

#### RESULT 5

LOCUS A2992478 29 bp DNA linear GSS 27-APR-2001  
DEFINITION 2M0277H07F Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0277H07 F, DNA sequence.

ACCESSION A2992478  
VERSION A2992478.1 GI:13863705

KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 29)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunne@genetics.utah.edu  
Insert length: 10000 Std Error: 0.00  
Plate: 0277 row: H column: 07  
Seq primer: CGTTGTAACAGCAGCCGACG  
Class: plasmid ends  
High quality sequence stop: 29.

FEATURES  
source 1. 29  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0277H07"  
/clone\_lib="Mouse 10kb plasmid UUGC2M library"  
/sex="Female"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource  
(<http://www.jax.org/resources/documents/dnares/>). The DNA



```

Plate: 0122  row: D  column: 06
Seq primer: CGTTGTAACACGACGGCCAGT
Class: plasmid  ends
High quality sequence stop: 21.
location/qualifiers

```

BASE COUNT	1 a	7 c	0 g	13 t
ORIGIN				

Query Match	100.0%	Score 11	DB 17	Length 21
Best Local Similarity	100.0%	Pred. NC	4e+04	
Matches 11	Conservative 0	Mismatches 0	Indels 0	Gaps 0

RESULT 2	LOCUS	DEFINITION
AZ861819	AZ861819	24 bp DNA linear GSS 21-FEB-2001
		2M0168112R Mouse 10kb plasmid uucgim library Mus musculus genomic clone uucg2M0168112 R, DNA sequence.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 24)	Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weiss, R.	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts	unpublished (2000)	Contact: Robert B. Weiss

```

FEATURES
  Plate: 0168  row: 1  column: 12
  Seq primer: CACACAGGAAACACAGCTATGACCT
  Class: plasmid ends
  High quality sequence stop: 24.
  Location/Qualifiers

```

BASE COUNT	9 a	0 c	15 g	0 t
ORIGIN				

Query Match	100.0%	Score 11;	DB 17;	Length 24;
Best Local Similarity	100.0%;	Pred. No.	4.1e+04;	
Matches 11; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

RESULT 3	BH789331/c	24 bp	DNA	linear	GSS 02-APR-2002
LOCUS	BH789331				
DEFINITION	SALK_019058.23.05.x Arabidopsis thaliana RNA insertion lines				
	Arabidopsis thaliana genomic clone SALK_019058.23.05.x, DNA				
	sequence.				

TITLE	A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL	Unpublished (2001)
COMMENT	Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (Signal)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: eckerssalk.edu

This is single pass sequence recovered from the left border of

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 09:24:34 ; Search time 21.0947 seconds  
(without alignments)  
8445.286 Million cell updates/sec

Title: US-09-902-772-1\_COPY\_650\_660

Perfect score: 11

Sequence: 1 agagagagagag 11

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inu:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	11	100.0	21	AZ371089	1M0122D06
c 2	11	100.0	24	AZ861819	AZ861819 2M0168L12
c 3	11	100.0	24	BH789331	SAUK_0190
c 4	11	100.0	25	AZ635993	1M0493E20
c 5	11	100.0	29	AZ992478	2M0277H07
c 6	11	100.0	33	AZ586653	AZ586653 1M0392P11

c 7	11	100.0	36	AZ351541	1M0089G11
c 8	11	100.0	37	BG744186	BG744186
c 9	11	100.0	38	BG684760	BG684760
c 10	11	100.0	40	AA922878	AA922878
c 11	11	100.0	40	BI252040	BI252040
c 12	11	100.0	46	BH812871	SAUK_0633
c 13	11	100.0	52	BF651086	BF651086
c 14	11	100.0	54	BM395183	50072-2-7
c 15	11	100.0	54	AZ996539	2M0282F14
c 16	11	100.0	56	AZ983943	2M0255G18
c 17	11	100.0	57	AZ345621	1M0080F19
c 18	11	100.0	59	AL767458	Arabidops
c 19	11	100.0	60	BQ767486	EBR008_SQ
c 20	11	100.0	61	AA503506	nh59h11.s
c 21	11	100.0	63	BH857098	SAUK_0768
c 22	11	100.0	68	BM521350	sa113g09.
c 23	11	100.0	68	AZ357658	1M0099F24
c 24	11	100.0	69	AL219577	Tetradodon
c 25	11	100.0	75	N85423	J3490F_Huma
c 26	11	100.0	75	AZ506280	1M0347B18
c 27	11	100.0	76	BQ094017	f54604.Y
c 28	11	100.0	76	AZ805134	2M0056G05
c 29	11	100.0	79	AA668503	ab86h11.s
c 30	11	100.0	79	AI082539	ox55f04.s
c 31	11	100.0	79	AI085281	qf18e06.x
c 32	11	100.0	79	AA142830	z140b01.s
c 33	11	100.0	79	AA620456	ae58d08.s
c 34	11	100.0	79	AZ471783	1M0286O22
c 35	11	100.0	79	AZ816527	2M0085M14
c 36	11	100.0	79	L8315BX	Leishman1
c 37	11	100.0	81	BM301739	MCA044A08
c 38	11	100.0	81	AO073887	EP(3)3288
c 39	11	100.0	84	BE324070	NFO13B03P
c 40	11	100.0	85	AI152770	u094f01.t
c 41	11	100.0	85	AI438348	SMOVARCAP
c 42	11	100.0	85	AI225000	1006123B0
c 43	11	100.0	85	AL181649	Tetradodon
c 44	11	100.0	87	BF718468	EST241.m1
c 45	11	100.0	88	AZ860469	AZ860469 2M0166A06

## ALIGNMENTS

RESULT 1  
LOCUS  
DEFINITION  
1M0122D06F Mouse 10kb plasmid tUGCIM library Mus musculus genomic  
clone tUGCIM0122D06 F, DNA sequence.

ACCESSION  
AZ371089  
KEYWORDS  
GSS.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus.  
REFERENCE  
AUTHORS  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.

TITLE  
Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 10:09:50 ; Search time 3.03292 Seconds  
(without alignments)  
7482.227 Million cell updates/sec

Title: US-09-902-772-1\_COPY\_650\_660

Perfect score: 11

Sequence: 1 agagagagagag 11

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCIT\_NEM\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEM\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEM\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCITUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEM\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEM\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEM\_PUB.seq:2:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:3:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:4:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEM\_PUB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEM\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	19	US-09-969-373-3925	Sequence 3925, Ap
2	11	100.0	20	US-10-215-112-9210	Sequence 9210, Ap
3	11	100.0	20	US-10-215-112-9325	Sequence 9325, Ap
4	11	100.0	20	US-10-215-112-9435	Sequence 9435, Ap
5	11	100.0	20	US-10-215-112-9437	Sequence 9437, Ap
6	11	100.0	20	US-10-215-112-9439	Sequence 9439, Ap
7	11	100.0	20	US-10-215-112-9715	Sequence 9715, Ap
8	11	100.0	20	US-10-215-112-9829	Sequence 9829, Ap
9	11	100.0	20	US-10-215-112-9938	Sequence 9938, Ap
10	11	100.0	20	US-10-215-112-9940	Sequence 9940, Ap
11	11	100.0	20	US-10-215-112-9942	Sequence 9942, Ap
12	11	100.0	24	US-09-852-980A-1	Sequence 1, Appl1
13	11	100.0	25	US-10-098-263B-10719	Sequence 10719, A
14	11	100.0	25	US-10-098-263B-10720	Sequence 10720, A
15	11	100.0	25	US-10-098-263B-22949	Sequence 22949, A
16	11	100.0	25	US-10-098-263B-63581	Sequence 63581, A

c	17	11	100.0	25	15	US-10-098-263B-80930	Sequence 80930, A
	18	11	100.0	25	15	US-10-098-263B-117135	Sequence 117135, A
	19	11	100.0	63	15	US-10-081-347-12	Sequence 12, Appl
	20	11	100.0	90	11	US-09-974-300-4057	Sequence 4057, Ap
	21	11	100.0	98	11	US-09-969-373-100	Sequence 100, Ap
	22	11	100.0	99	10	US-09-864-761-18548	Sequence 18548, A
	23	11	100.0	104	11	US-09-783-590-6469	Sequence 6469, A
	24	11	100.0	111	10	US-09-864-761-25653	Sequence 25653, A
	25	11	100.0	117	10	US-09-864-761-32279	Sequence 32279, A
c	26	11	100.0	120	11	US-09-969-373-1197	Sequence 1197, A
	27	11	100.0	124	10	US-09-864-761-17723	Sequence 17723, A
	28	11	100.0	129	11	US-09-728-444-210	Sequence 210, Appl
	29	11	100.0	141	15	US-10-081-347-19	Sequence 19, Appl
	30	11	100.0	146	11	US-09-728-444-211	Sequence 211, Appl
	31	11	100.0	150	10	US-09-728-446-882	Sequence 882, Appl
c	32	11	100.0	156	10	US-09-923-876-5331	Sequence 5331, Appl
	33	11	100.0	162	11	US-09-878-574-11579	Sequence 11579, A
c	34	11	100.0	165	11	US-09-878-574-9778	Sequence 9778, Ap
	35	11	100.0	170	10	US-09-864-761-25721	Sequence 25721, A
c	36	11	100.0	171	10	US-09-728-446-882	Sequence 882, Appl
	37	11	100.0	174	11	US-09-764-847-1715	Sequence 1715, Appl
c	38	11	100.0	174	15	US-10-092-154-1715	Sequence 1715, Appl
c	39	11	100.0	176	11	US-09-764-847-1717	Sequence 1717, Ap
c	40	11	100.0	176	15	US-10-092-154-1717	Sequence 1717, Ap
	41	11	100.0	177	10	US-09-770-696-299	Sequence 299, Appl
c	42	11	100.0	178	10	US-09-864-761-25409	Sequence 25409, A
	43	11	100.0	180	10	US-09-294-093B-3061	Sequence 3061, Appl
c	44	11	100.0	182	15	US-10-085-906-159	Sequence 159, Appl
	45	11	100.0	183	11	US-09-867-701-8586	Sequence 8586, Appl

## ALIGNMENTS

RESULT 1  
US-09-969-373-3925  
Sequence 3925, Application US/09969373  
Patent No. US20020133852A1  
GENERAL INFORMATION:  
APPLICANT: Efferitz, Roger J.  
TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping  
FILE REFERENCE: 38-10(52679)A  
CURRENT APPLICATION NUMBER: US/09/969,373  
CURRENT FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: US 09/754,853  
PRIOR FILING DATE: 2001-01-05  
PRIOR APPLICATION NUMBER: US 09/760,427  
PRIOR FILING DATE: 2001-01-13  
PRIOR APPLICATION NUMBER: US 09/855,768  
PRIOR FILING DATE: 2001-09-15  
NUMBER OF SEQ ID NOS: 4593  
SEQ ID NO 3925  
LENGTH: 19  
TYPE: DNA  
ORGANISM: Glycine max  
US-09-969-373-3925

Query Match 100.0%; Score 11; DB 11; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.7e+03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 AGAGAGAGGAG 11  
1 AGAGAGAGGAG 11  
|||||

RESULT 2  
US-10-215-112-9210  
Sequence 9210, Application US/10215112  
Publication No. US20030082596A1  
GENERAL INFORMATION:  
APPLICANT: Michael Miltmann

```

; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE OF INVENTION: Test3
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 9210
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9210

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 3
US-10-215-112-9325
; Sequence 9325, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 9325
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9325

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 4
US-10-215-112-9435
; Sequence 9435, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 9435
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9435
```

```

US-10-215-112-9435

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 5
US-10-215-112-9437
; Sequence 9437, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 9437
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9437

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 6
US-10-215-112-9439
; Sequence 9439, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 9439
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9439

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 7
US-10-215-112-9715/c
```

```
; Sequence 9715, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9715
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9715

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGGAG 11
        |||
Db      15 AGAGAGAGGAG 5

RESULT 8
US-10-215-112-9829/c
; Sequence 9829, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9829
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9829

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGGAG 11
        |||
Db      15 AGAGAGAGGAG 5

RESULT 9
US-10-215-112-9938/c
; Sequence 9938, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9938
; LENGTH: 20
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9938

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGGAG 11
        |||
Db      15 AGAGAGAGGAG 5

RESULT 10
US-10-215-112-9940/c
; Sequence 9940, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9940
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9940

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGGAG 11
        |||
Db      15 AGAGAGAGGAG 5

RESULT 11
US-10-215-112-9942/c
; Sequence 9942, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9942
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9942

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGGAG 11
        |||
Db      15 AGAGAGAGGAG 5
```

```
RESULT 12
US-09-852-980A-1
; Sequence 1, Application US/09852980A
; Patent No. US20020034752A1
; GENERAL INFORMATION:
; APPLICANT: ORDOVAS, JOSE M.
; APPLICANT: SCHAEFER, ERNST J.
; TITLE OF INVENTION: CPT TACIB POLYMORPHISM AS RISK FACTOR FOR DEVELOPMENT
; TITLE OF INVENTION: OF CORONARY HEART DISEASE
; FILE REFERENCE: TU-2004
; CURRENT APPLICATION NUMBER: US/09/852,980A
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,467
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-852-980A-1

Query Match
Best Local Similarity 100.0%; Score 11; DB 10; Length 24;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 10 AGAGAGAGGAG 20

RESULT 13
US-10-098-263B-10719
; Sequence 10719, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 10719
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-10719

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 1 AGAGAGAGGAG 11

RESULT 14
US-10-098-263B-10720
; Sequence 10720, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
```

```
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 10720
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-10720

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 1 AGAGAGAGGAG 11

RESULT 15
US-10-098-263B-22949
; Sequence 22949, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 22949
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-22949

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 8 AGAGAGAGGAG 18
```

Search completed: July 28, 2003, 12:23:26  
Job time: 5.03292 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 07:39:19 ; Search time 4232.82 Seconds

(without alignments)  
\$948.861 Million cell updates/sec

Title: US-09-902-772-1

Perfect score: 1447

Sequence: 1 gaattccggaacgaataat.....gaaagaagcgccaagaanaa 1447

Scoring table: .IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml:\*

1: gb\_ba:\*

2: gb\_hlg:\*

3: gb\_in:\*

4: gb\_com:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_dr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_com:\*

21: em\_ov:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_hlg\_hum:\*

31: em\_hlg\_inv:\*

32: em\_hlg\_other:\*

33: em\_hlg\_mus:\*

34: em\_hlg\_pln:\*

35: em\_hlg\_rtd:\*

36: em\_hlg\_mam:\*

37: em\_hlg\_vrt:\*

38: em\_sy:\*

39: em\_higo\_hum:\*

40: em\_higo\_mus:\*

41: em\_higo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1447	100.0	1447	6 E31253	E31253 Protein hav
2	1409	97.4	1413	5 A1055661	A1055661 Gallus ga
3	1356	93.7	1528	6 E31254	E31254 Protein hav
4	1344	92.9	1516	5 GGERG	X77129 G.gallus ER
5	936	64.7	1808	10 AB031088	AB031088 Rattus no
6	921	63.6	3166	9 HUMERG2	M17254 Human erg2
7	889.6	61.5	2012	5 XLAJ4126	AJ224126 Xenopus 1
8	881.6	60.9	2131	10 AB073078	AB073078 Mus muscu
9	867.2	59.9	2209	10 AB073080	AB073080 Mus muscu
10	828	57.2	2133	10 AB073079	AB073079 Mus muscu
11	801.8	55.4	1798	5 XLAJ4125	AJ224125 Xenopus 1
12	760.2	52.5	3126	9 HUMERG11	M21535 Human erg p
13	571.2	39.5	1436	5 DRE249590	AJ249590 Danio rer
14	570.2	39.4	1729	10 MPEL11	X59421 Mouse Flt-1
15	563.8	39.0	2916	6 A36461	A36461 Sequence 2
16	563.8	39.0	2938	6 AR080101	AR080101 Sequence
17	563.8	39.0	2938	6 HSHMFL1	X67001 H.sapiens H
18	562.2	38.9	1359	9 AY029368	AY029368 Homo sapi
19	562.2	38.9	2394	9 BC010115	BC010115 Homo sapi
20	562.2	38.9	2957	9 HUMERGBFL1	M98833 Homo sapien
21	562.2	38.9	2959	9 BC001670	BC001670 Homo sapi
22	561.4	38.8	6849	12 AY029367	AY029367 Synthetic
23	560.6	38.7	1673	9 S45205	S45205 Flt-1-Frien
24	549.4	38.0	4403	5 XLF11G	X66979 X.laevis nr
25	518.6	35.8	1245	5 AF177538	AF177538 Danio rer
26	514	35.5	3450	5 CCNAPFL1	Y14773 Colurnix co
27	502.2	34.7	3545	5 CCEFLIONCO	Y14774 Colurnix co
28	493.2	34.1	762	9 S72621	S72621 EMS...erg f
29	478.2	33.0	1932	9 HUMFL1A	M93255 Human Flt-1
30	477	33.0	816	10 S66169	S66169 Erg-3-immun
31	387	26.7	14518	9 AP001423	AP001423 Homo sapi
32	387	26.7	100000	9 AP000021	AP000021 Homo sapi
33	387	26.7	100000	9 AP000163	AP000163 Homo sapi
34	387	26.7	340000	9 AP001731	AP001731 Homo sapi
35	373	25.8	588	9 S72865	S72865 EMS...EMS-e
36	358.6	24.8	1431	9 AF327066	AF327066 Homo sapi
37	320.4	22.1	66352	9 HSY17293	Y17293 Homo sapien
38	320.4	22.1	187131	2 AP001535	AP001535 Homo sapi
39	320.4	22.1	217382	9 AP001122	AP001122 Homo sapi
40	214.4	14.8	634	3 SULERG	M81067 Sea urchin
41	203	14.0	491	3 HD1311813	AJ311813 Hediste d
42	196.6	13.6	2180	3 AY060316	AY060316 Drosophila
43	180	12.4	1403	9 HSDNAFEV3	Y08979 H.sapiens F
44	180	12.4	1752	10 RN091679	Y091679 Rattus norv
45	180	12.4	169741	9 AC097468	AC097468 Homo sapi

#### ALIGNMENTS

RESULT 1

E31253

LOCUS E31253 1447 bp DNA linear PAT 18-JUN-2001

DEFINITION Protein having cell calcifying inhibitory activity and gene

ACCESSION E31253

VERSION E31253.1 GI:13025685

KEYWORDS JP 1999075871-A/1.

SOURCE unidentifed.

ORGANISM unidentifed.

REFERENCE 1 (bases 1 to 1447)

AUTHORS Hiroyasu,I., Yoshinobu,H., Marjio,P., Joel,R. and Helena,E.

TITLE Protein having cell calcifying inhibitory activity and gene encoding the same

Pred. No. is the number of results predicted by chance to have a





361 GAAAGAGAGCAATATCCGCCCTCCAAATATGACACACCAACCAATGAAACGAGAGTTATGTCCCA 420

LOCUS	DEFINITION	Protein having cell calcifying inhibitory activity and gene
Db	447	GCAGATCTCTACGTTATGGACGACAGACCATCTACGGAGTGGCTGGAGTGGCCAGTAAG 508
Db	421	GCAGATCTCTACGTTATGGAGCAGACCATCTATACGGCAGTGGCTGGAGTGGCCAGTAAG 480
Qy	507	GAGTATGCTCTCCAGAGCTGGACATCTGTTGTTCCGAAACATTTGATGGGAAGAGTGG 566
Db	481	GAGTATGCTCTCCAGAGCTGGACATCTGTTGTTCCGAAACATTTGATGGGAAGAGTGG 540
Qy	567	TGTAAATAGACCAAAAGATGACTTCCAGAGACTACAGCCGAGCTATAACGCAGATATCCTC 626
Db	541	TGTAAATAGACCAAAAGATGACTTCCAGAGACTACAGCCGAGCTATAACGCAGATATCCTC 600
Qy	627	CTGTACACCTTACACTACCTCAGAGAGAGAGAGGACCACTTTATTTTCCAAATACATCA 686
Db	601	CTGTACACCTTACACTACCTCAGAGAGAGAGAGGACCACTTTATTTTCCAAATACATCA 660
Qy	687	GTTTATCCCAAGCAAGCAAGCAAGATATACCAAGGCGAGATTTATACCTTATGAGCAAGG 746
Db	661	GTTTATCCCAAGCAAGCAAGCAAGATATACCAAGGCGAGATTTATACCTTATGAGCAAGG 720
Qy	747	AGAGATACAGCGTGGAGAGTACAGAGCCATCCCACTAGTCAAAAGCTATACCAACATCA 806
Db	721	AGAGATACAGCGTGGAGAGTACAGAGCCATCCCACTAGTCAAAAGCTATACCAACATCA 780
Qy	807	TCTTCAACAGTGCCTCAAAACAGAAAGACCAGCGTCTCAGTTAGATTCCTTATCAGATTCCT 866
Db	781	TCTTCAACAGTGCCTCAAAACAGAAAGACCAGCGTCTCAGTTAGATTCCTTATCAGATTCCT 840
Qy	887	GGACCGACACAGCGCTGCTGCAATCCAGGAGTGGCGAGATACAGATATGGCACTTC 926
Db	841	GGACCGACACAGCGCTGCTGCAATCCAGGAGTGGCGAGATACAGATATGGCACTTC 900
Qy	927	CTACTGAGCTCTTGTGTGGAGAGTCCCAACTCCACATGACACTCTGGAGGCGACAAAT 986
Db	901	CTACTGAGCTCTTGTGTGGAGAGTCCCAACTCCACATGACACTCTGGAGGCGACAAAT 960
Qy	987	GGGAGATTCAAGATGACAGACCCGTGATGAGTGGCTGGGTTGGGAGAGAGGAAAGC 1046
Db	961	GGGAGATTCAAGATGACAGACCCGTGATGAGTGGCTGGGTTGGGAGAGAGGAAAGC 1020
Qy	1047	AAACCTACATGAAGTATGACAAATCAAGCCGTGCTACCTGCTACTATGACAAAT 1106
Db	1021	AAACCTACATGAAGTATGACAAATCAAGCCGTGCTACCTGCTACTATGACAAAT 1080
Qy	1107	ATTATGACTAAAGTTTCAATGTTAAACGCTATGCTTCAAAATTTGATTTCCAGGAATGCT 1166
Db	1081	ATTATGACTAAAGTTTCAATGTTAAACGCTATGCTTCAAAATTTGATTTCCAGGAATGCT 1140
Qy	1167	CAGGCCCTCAGCCCTCACCTCCAGAAATCATCATGTAACAATACCATCAGACCTCCC 1226
Db	1141	CAGGCCCTCAGCCCTCACCTCCAGAAATCATCATGTAACAATACCATCAGACCTCCC 1200
Qy	1227	TACATGAGTTCCTACATGACACACCCCAAGAAATGAATTTGATGCTCCCCATCCCTC 1286
Db	1201	TACATGAGTTCCTACATGACACACCCCAAGAAATGAATTTGATGCTCCCCATCCCTC 1260
Qy	1287	GCTTGGCCCGTAACCTCATCCAGCTTTTTCCTGCTCCCTTAATCATACTGGAATTCACCA 1346
Db	1261	GCTTGGCCCGTAACCTCATCCAGCTTTTTCCTGCTCCCTTAATCATACTGGAATTCACCA 1320
Qy	1347	ACTGGAGGCATCTACCCCAATATCCAGGCTGGCAGCTCTCATATGCTCTCCATCTTGGC 1406
Db	1321	ACTGGAGGCATCTACCCCAATATCCAGGCTGGCAGCTCTCATATGCTCTCCATCTTGGC 1380
Qy	1407	ACCTACTACTATAGTGGGGAAGAAAGAAA 1435
Db	1381	ACCTACTACTATAGTGGGGAAGAAAGAAA 1409

encoding the same.  
 accession E31254  
 version E31254.1 GI:13025686  
 keywords JP 1999075871-A/2.  
 source unidentified.  
 organism unidentified.  
 reference 1 (bases 1 to 1528)  
 authors Hiroyasu, I., Yoshinobu, H., Marijo, P., Joel, R. and Helena, E.  
 title Protein having cell calcifying inhibitory activity and gene  
 encoding the same  
 patent: JP 1999075871-A 2 23-MAR-1999;  
 CHUGAI PHARMACEUT CO LTD, HE TRUSTEES OF THE UNIVERSITY OF  
 PENNSYLVANIA  
 journal

COMMENT  
 OS Unidentified  
 PN JP 1999075871-A/2  
 PD 23-MAR-1999  
 PF 29-MAY-1998 JP 1998166076  
 PR 18-JUN-1997 US 08/878177, 20-JUN-1997 US 60/050297 PI  
 HIROYASU IMAMOTO, YOSHINOBU HIGUCHI, MARIJO PASHIRIKI, PI JOEL  
 ROZENBURG, PI HELENA E

PC C12N15/09, A61K48/00, C07K14/465, C07K16/18, C12Q1/68//A61K38/00,  
 PC C12P21/02,  
 PC C12P21/08, C12N15/00, A61K37/02  
 CC Strandedness: Single;  
 CC Topology: Linear;  
 CC Key  
 FT source  
 FT Location/Qualifiers  
 1..1528  
 /organism="Unidentified",  
 /db\_xref="taxon:32644"

BASE COUNT 464 a 395 c 330 g 339 t  
 ORIGIN

Query Match 93.7%; Score 1356; DB 6; Length 1528;  
 Best Local Similarity 94.7%; Pred. No. 0;  
 Matches 1447; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 1 GAATTCGGCAAGCAATATTTATTATTTAGCAATTTATGCGATCAATATCTTGATCACA 60  
 DB 1 GAATTCGGCAAGCAATATTTATTATTTAGCAATTTATGCGATCAATATCTTGATCACA 60  
 QY 61 TTATGGCAACACTATTATTAAGCAATTTATGAGTGTGAGTGAAGACAGCTCTGTTG 120  
 DB 61 TTATGGCAACACTATTATTAAGCAATTTATGAGTGTGAGTGAAGACAGCTCTGTTG 120  
 QY 121 AGTGTGCTACGATCGCCCTGCAAGACAGAAATGACAGCTCTCTTCAGTG 180  
 DB 121 AGTGTGCTACGATCGCCCTGCAAGACAGAAATGACAGCTCTCTTCAGTG 180  
 QY 181 AATATGGGCAAAATCAATCAATATGAGCCCGGCTTCCCGACAGAGTGTATACAGC 240  
 DB 181 AATATGGGCAAAATCAATCAATATGAGCCCGGCTTCCCGACAGAGTGTATACAGC 240  
 QY 241 CCCCAGCAGAGTATACATTAAGATGAGTGTAAACCAACAGAGTGTATGAGTCAAGA 300  
 DB 241 CCCCAGCAGAGTATACATTAAGATGAGTGTAAACCAACAGAGTGTATGAGTCAAGA 300  
 QY 301 ATTACCTGATGACTGACGCTGGCAAAAGAGGAGAAATGTTAGCAGTTCAAGC 360  
 DB 301 ATTACCTGATGACTGACGCTGGCAAAAGAGGAGAAATGTTAGCAGTTCAAGC 360  
 QY 361 TTGGGATGAACTATGAGAGTACATGAGAGAGAGATATTCGCGCTCCAAATATGACA 420  
 DB 361 TTGGGATGAACTATGAGAGTACATGAGAGAGAGATATTCGCGCTCCAAATATGACA 420  
 QY 421 CCAATGAAAGAGAGTATGAGTGTGAGAGATCTAGTATGAGAGAGAGAGATGATAC 480  
 DB 421 CCAATGAAAGAGAGTATGAGTGTGAGAGATCTAGTATGAGAGAGAGAGATGATAC 480

QY 481 GCGAGTGGCTGAGTGGGAGAGAGATGATGCTTCCAGACGTGAGATCTTGTGT 540  
 DB 481 GCGAGTGGCTGAGTGGGAGAGATGATGCTTCCAGACGTGAGATCTTGTGT 540  
 QY 541 TCCAGAACTATGATGGGAAAGAGTGTGTAAATGACCAAAAGATGACTTCAGAGACTCA 600  
 DB 541 TCCAGAACTATGATGGGAAAGAGTGTGTAAATGACCAAAAGATGACTTCAGAGACTCA 600  
 QY 601 CCGCGAGCTTAAACGAGATATCTCTGTCACACCTACCTACCTCTCAGAGACA ----- 654  
 DB 601 CCGCGAGCTTAAACGAGATATCTCTGTCACACCTACCTACCTCTCAGAGACTCTC 660  
 QY 655 ----- 654  
 DB 661 TTCACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 QY 655 -----GAGAGCCACTTTATTTTCCAAATATCATGTTTACCCAGAAAG 699  
 DB 721 ATCTAGAAACACAG 780  
 QY 700 CAACGCAAGATTAACAAAG 759  
 DB 781 CAACGCAAGATTAACAAAG 840  
 QY 760 GAGAGAGTACAG 819  
 DB 841 GAGAGAGTACAG 900  
 QY 820 CCAAAACAGAAAG 879  
 DB 901 CCAAAACAGAAAG 960  
 QY 880 GCGGTCTTGAATTCAG 939  
 DB 961 GCGGTCTTGAATTCAG 1020  
 QY 940 TGTGGAGAGCTCCAACTCCAACTGATCATCTGGAGAGAGAGAGAGAGAGAGAGAG 999  
 DB 1021 TGTGGAGAGCTCCAACTCCAACTGATCATCTGGAGAGAGAGAGAGAGAGAGAGAG 1080  
 QY 1000 TGACAGAGCTGATGAGAGTGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1059  
 DB 1081 TGACAGAGCTGATGAGAGTGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
 QY 1060 ACTATGCAAACTCAG 1119  
 DB 1141 ACTATGCAAACTCAG 1200  
 QY 1120 TTCATGTAAGAGCTATGCTTACAAATTTGATTTTCAGAGAGAGAGAGAGAGAGAG 1179  
 DB 1201 TTCATGTAAGAGCTATGCTTACAAATTTGATTTTCAGAGAGAGAGAGAGAGAGAG 1260  
 QY 1180 CTCACCTCCAGAGATATCATGATTAACCAATACCAATACCAATACCAATACCAATAC 1239  
 DB 1261 CTCACCTCCAGAGATATCATGATTAACCAATACCAATACCAATACCAATACCAATAC 1320  
 QY 1240 ACCATGCAACCCCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1299  
 DB 1321 ACCATGCAACCCCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
 QY 1300 CCTCATCAGCTTTTGTGCTGCTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 1359  
 DB 1381 CCTCATCAGCTTTTGTGCTGCTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 1440  
 QY 1360 ACCGCAATCCAGAGTGGAG 1419  
 DB 1441 ACCGCAATCCAGAGTGGAG 1500  
 QY 1420 TGGGGAAG 1447  
 DB 1501 TGGGGAAG 1528

RESULT 4  
 GSERG LOCUS 1516 bp mRNA linear VRT 27-Apr-1995  
 DEFINITION G.gallus ERG mRNA.  
 X77159  
 VERSION X77159.1 GI:790439  
 KEYWORDS ERG gene.  
 SOURCE chicken.  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 1516)  
 Dideric, P., Dewitte, F., Desbiens, X., Stehelin, D. and  
 Duterque-Coquillaud, M.  
 Mesodermal expression of the chicken erg gene associated with  
 precartilaginous condensation and cartilage differentiation  
 Mech. Dev. 50 (1), 17-28 (1995)  
 95329425  
 7605748  
 2 (bases 1 to 1516)  
 Duterque-Coquillaud, M.  
 Direct Submission  
 Submitted (17-JAN-1994) M. Duterque-Coquillaud, CNRS UA 1160,  
 Oncologie Moléculaire, Institut Pasteur, 1 rue Calmette, 59019  
 Lille, FRANCE  
 FEATURES  
 source location/Qualifiers  
 1..1516  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"  
 /tissue\_type="spleen"  
 /clone\_id="lambda gt10"  
 /dev\_stage="adult"  
 /gene="ERG"  
 /gene="ERG"  
 63..1499  
 /gene="ERG"  
 /codon\_start=1  
 /protein\_id="CAA54404.1"  
 /db\_xref="GI:790440"  
 /translation="MASTIKELALVSEVSEDOSELECAVSGPFLAKTEMTASSSEYQGT  
 SKSPRPQDDMLSOPARYTIKMECPNPNVNSPDCEVAKGKWMSSSDNGM  
 NGSYMEKHIPPNMTNERIVIPADPTLMSTDHROWLHNAVVEYGLPDVILLF  
 QNIDKELCKMTKDDDFQRLPSYNADILLSHLRLKTRPLPHLTSDDVYKALONSRL  
 MHRNKGATFIFPNTSVYPEATQRTTRDLPYEQRRSAMTSHSHPTQSKATQBS  
 STVKTEDDPPQILGPTSSRLNPSGQIOQFLEILSDSSNRCITWGT  
 NGFEKMTDPDEVARMRGERSKPMNMDKLSRLRYDDKINIMTKVHGKRYAKPFEH  
 GIAOALOPHESSMYKYPEDLPYMSYHHPQKMFVAPHPALPYTSSFFAAPRP  
 YMNSPTGTYPRNTRIPAHMPSHLGTYT"  
 BASE COUNT 458 a 392 c 327 g 339 t  
 ORIGIN  
 Query Match 92.9%; Score 1344; DB 5; Length 1516;  
 Best Local Similarity 94.7%; Pred. No. 0;  
 Matches 1435; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 241 CCCGCGCAGAGTTACCATTAAGATGAGTGTAAACCAACAGAGTTATGGTCAAGA 300  
 |||||  
 Db 241 CCCCCGCGCAGAGTTACCATTAAGATGAGTGTAAACCAACAGAGTTATGGTCAAGA 300  
 QY 301 ATTACCTGATGACTGCGAGCGTGGCAAAAGAGGAAATGTTAGCAGTTCAACAATG 360  
 |||||  
 Db 301 ATTACCTGATGACTGCGAGCGTGGCAAAAGAGGAAATGTTAGCAGTTCAACAATG 360  
 QY 361 TTGGGATGAACCTATGAGAGCTACATGGAAGAAGCAATTCGCCCTCCAAATATGCAA 420  
 |||||  
 Db 361 TTGGGATGAACCTATGAGAGCTACATGGAAGAAGCAATTCGCCCTCCAAATATGCAA 420  
 QY 421 CCAATGAACAGAGATGATTTGTGCACAGATCCTACGTTATGAGCAGACCATGTAC 480  
 |||||  
 Db 421 CCAATGAACAGAGATGATTTGTGCACAGATCCTACGTTATGAGCAGACCATGTAC 480  
 QY 481 GGCAGTGGCTGAGTGGGAGTGAAGAGATGATGTTCCAGAGCTGGACATCTTGTGT 540  
 |||||  
 Db 481 GGCAGTGGCTGAGTGGGAGTGAAGAGATGATGTTCCAGAGCTGGACATCTTGTGT 540  
 QY 541 TCCAGAACATTTGATGGGAAAGAGTTGTCTAAATGACCAAAAGATGACTTCCAGAGCTCA 600  
 |||||  
 Db 541 TCCAGAACATTTGATGGGAAAGAGTTGTCTAAATGACCAAAAGATGACTTCCAGAGCTCA 600  
 QY 601 CGCGAGCTATTAAGCAGATATGCTCCTGTCACACCTACACTACCTCAGAGAGA----- 654  
 |||||  
 Db 601 CGCGAGCTATTAAGCAGATATGCTCCTGTCACACCTACACTACCTCAGAGAGATCTCTC 660  
 QY 655 ----- 654  
 |||||  
 Db 661 TTCCACATTTGACTTCAGATGATGTTGATTAAGGCTTACAAACCTCCACGGTTAATGC 720  
 QY 655 -----GAGGAGCCACTTTATTTTTCAAATATACATGTTACCAGAG 659  
 |||||  
 Db 721 ATGCTAGAAACACAGAGGAGGAGCACTTTATTTTCCAAATATACATGTTACCAGAG 780  
 QY 700 CAACGCAAAAGATTAACAACAAGGAGCAGATTACTTGTGAGCAAGCAGAGATCAGCGT 759  
 |||||  
 Db 781 CAACGCAAAAGATTAACAACAAGGAGCAGATTACTTGTGAGCAAGCAGAGATCAGCGT 840  
 QY 760 GAGAGATCAGAGCATCCAGCTCAGTCAAAAGCTTACCAACATCTTCAACAGTGC 819  
 |||||  
 Db 841 GAGAGATCAGAGCATCCAGCTCAGTCAAAAGCTTACCAACATCTTCAACAGTGC 900  
 QY 820 CCAAAACAGAGACACCGCTCTCAGTTAGATCCTTATCGATTTTGGACCAACAGCA 879  
 |||||  
 Db 901 CCAAAACAGAGACACCGCTCTCAGTTAGATCCTTATCGATTTTGGACCAACAGCA 960  
 QY 880 GCGGCTTGGCAATCCAGGAGAGTGGGAGTACAGTATGAGCAGTTCCTCTGAGAGTTC 939  
 |||||  
 Db 961 GCGGCTTGGCAATCCAGGAGAGTGGGAGTACAGTATGAGCAGTTCCTCTGAGAGTTC 1020  
 QY 940 TGTGCGACAGCTCCAACTCCAACTGCATCCAGTGGGAGGACACAAATGGGAGTTCAGA 999  
 |||||  
 Db 1021 TGTGCGACAGCTCCAACTCCAACTGCATCCAGTGGGAGGACACAAATGGGAGTTCAGA 1080  
 QY 1000 TGACAGACCTGATGAAGTGGCTCGCGGTTGGGAGAGAGAGAAAGCAAACTTAACATGA 1059  
 |||||  
 Db 1081 TGACAGACCTGATGAAGTGGCTCGCGGTTGGGAGAGAGAGAAAGCAAACTTAACATGA 1140  
 QY 1060 ACTATGACAACCTCAGCGCTGCACTTGCCTACTACTATGACAAAAATATATATGAAG 1119  
 |||||  
 Db 1141 ACTATGACAACCTCAGCGCTGCACTTGCCTACTACTATGACAAAAATATATATGAAG 1200  
 QY 1120 TTCTAGGTAAAGCTATGCTTACAAATTTGATTTCCAGGAAATCGCTCAGGCCCTCAGC 1179  
 |||||  
 Db 1201 TTCTAGGTAAAGCTATGCTTACAAATTTGATTTCCAGGAAATCGCTCAGGCCCTCAGC 1260  
 QY 1180 CTCACCTCCAGATATCATCATGTACAAATACCATCAGACCTCCCTACATGATTTCT 1239  
 |||||  
 Db 1261 CTCACCTCCAGATATCATCATGTACAAATACCATCAGACCTCCCTACATGATTTCT 1320  
 QY 1240 ACCATGACACCCCGCAGAAAGATGAGTGTGTAGCTCCCATTCGCCCTGCTTCCCGGTAA 1299

```

|||||
Db 1321 ACATGACACACCCCAAGATGACCTTGTAGCTCCCATCCCGCTGCTTGGCCCTAA 1380
QY 1300 CCTCATCCAGCTTTTGTGCCCCATATCCATATGGAATTCACAAAGGAGGATCT 1359
Db 1381 CCTCATCCAGCTTTTGTGCCCCATATCCATATGGAATTCACAAAGGAGGATCT 1440
QY 1360 ACCCAATACCAAGCTCCAGCTGCTCATATGCCCTCCATCTTGACACCTACTACTAG 1419
Db 1441 ACCCAATACCAAGCTCCAGCTGCTCATATGCCCTCCATCTTGACACCTACTACTAG 1500
QY 1420 TGGGGAAGAAAGAA 1435
Db 1501 TGGGGAAGAAAGAA 1516

RESULT 5
AB031088 1808 bp mRNA linear ROD 07-AUG-2001
LOCUS DEFINITION
AB031088 Rattus norvegicus VESPL4 mRNA for vascular endothelial cell
AB031088 specific protein 14, complete cds.
VERSION AB031088.1 GI:15128488
KEYWORDS vascular endothelial cell specific protein 14.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine;
          Rattus.
REFERENCE 1 (bases 1 to 1808)
AUTHORS Aoki,T., Toyoda,H., Nishimoto,S., Tawara,J., Ukal,Y. and
          Komurasaki,T.
TITLE Identification of VESPL4, a vascular endothelial cell specific
          protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1808)
AUTHORS Aoki,T., Toyoda,H., Nishimoto,S., Tawara,J., Ukal,Y. and
          Komurasaki,T.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1999) Tomomi Aoki, Taisho Pharmaceutical Co.,
          Ltd., Medical Research Laboratories, Molecular Biology Lab.,
          1-403, Yoshino-cho, Omiya, Saitama 3308530, Japan
          (E-mail:s17704@com.taisho.co.jp, tel:048-663-1111,
          Fax:048-652-7254)

FEATURES
source location/Qualifiers
          1..1808
          /organism="Rattus norvegicus"
          /db_xref="taxon:10116"
          /tissue_type="liver"
          1..1808
          /gene="VESPL4"
          /gene="VESPL4"
          /note="similar to human ergl"
          /product="vascular endothelial cell specific protein 14"
          /protein_id="BAB62744.1"
          /translation="MASTRIKELMSVSKDLSFECAVGPPLAKTEMTASSSSDYGT
          SKSPVPOODMSOPPARVITKECNPSVNSRSPDECSITKGGKMGSPDTGM
          STGSIWEKHPPPNMTNRRVIVPADPLWSTDVHROWLENAVERIGLIDVLE
          ONIDGELCKMTKDDFORLPSYNADILSHLAYLRETPLPHLTSDVDALQNSPL
          MAHRNDLPEPERSRWGSHPTPOSKAQSPSTVEPTEDQRLDPYQILGPT
          SRLANGSGOIQLOMFLLELSDSNSNCTTEGTNGEETMDPDEVARWGBRKSQ
          NMNYDKLSRALRYVDKNIMTKYHGKRAYAKFDFHGIQAOLQHPRESSIXKPSDLP
          YMGSYTHPQKMFVAPHRPALPVYSSSPATPNPYWNSFTGSIYFNTRLPASHMFSH
          LGTY"
BASE COUNT 510 a 522 c 426 g 350 t
ORIGIN
Query Match 64.7%; Score 936; DB 10; Length 1808;
Best Local Similarity 80.2%; Fragment No. 5.8e-273;
Matches 1140; Conservative 0; Mismatches 270; Indels 12; Gaps 3;

```

```

QY 8 GCGAAGCAATATATATATAGCAATATATAGCATCAATATATCTGATCATATATGCG 67
Db 8 GAGAGCAATATATATATATATATATATATATATATATATATATATATATATATGCG 67
QY 68 AAGCATTATTAAGAGACATTATCAGTGTGATGAGAACACAGCTCTGTTGAGTGC 127
Db 68 CAGCATATTAAGAGACCTTGTGCTGATGAGAACAGACAGCTCATATTAAGTGTGC 127
QY 128 CTACGATGCCCCCACCCTTGCCAAAGACAGAAATGACAGCTCTCTCCAGTAATATG 187
Db 128 CTACGAGAGCCACACCTTGCCCTAAGACAGAAATGACAGCTCTCTCCAGTACATG 187
QY 188 GCAAAATCAAGATGAGCCCGGCTTCCACAGACAGCTGTTATACAGCCCGCG 247
Db 188 CCAAGATCCCAAGATGATGATCCAGAGTCCCAAGACAGAGCTGCTGATCAGCCCGAG 247
QY 248 CAGAGTTACATTATAGATGAGTGAATCCAAACAGCTTAATATGCTCAAGAAATTCAC 307
Db 248 CAGGCTCACCATCAAGATGAGTGAATCCCAAGCTTACAGCTGATACGTTCCAGAACTCAC 307
QY 308 TGATGACTGAGGCTGCGCAAGAGGGAATGATGATGATGATGATGATGATGATGAT 367
Db 308 TGAAGATGAGTGTGACCAAGGTGAGATGATGATGATGATGATGATGATGATGAT 367
QY 368 GAACATGAGAGCTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 427
Db 368 GAGCTAGAGAGCTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 427
QY 428 ACGAAGATTAATGTCGACAGACATCTTATGAGAGACAGACAGATGACGAGTG 487
Db 428 ACCGAGATGATGTCCTGTCAGATCTTCTGTGAGAGACAGACAGATGTCGGAGTG 487
QY 488 GCTGAGTGGGAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 547
Db 488 GCTGAGTGGGAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 547
QY 548 CATTTGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607
Db 548 TATTTGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607
QY 608 CTATTAAGCAGATATCTCTCTGTCACACCTTACCTTACCTTACCTTACCTTACCTT 662
Db 608 CTACAAATGCGAGATCTCTCTGTCACACCTTACCTTACCTTACCTTACCTTACCTT 662
QY 663 ACTTTATTTTCCAAATACATCAGTTTACCCAGAA-----GCAACGCAAGATTAACAC 718
Db 668 TCTGACATCCGACAGAGCTTATATAGGCTTTACAAACTCTCCACGGTTAATGATGCTAG 727
QY 719 AAGGCCAGATTTACCTTATGAGCAGACGAGAGATCAGCTGAGAGATCAGACGATCC 778
Db 728 AAGCAGATTTACCTTATGAGCCTCCACAGAGATCAACCTGAGACGGCCAGACGATCC 787
QY 779 CA---CTCAGTCAAAAGCTTACCCACCATATTTTCAACAGTCCCAAAACAGAACCA 835
Db 788 CACCCCTCATCCAAAGCTTCCACGAGATCTCTCCACAGTCCCAAAACAGAACCA 847
QY 836 GCGTCTCAGTTGATCCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 895
Db 848 GCGTCTCAGTTGATCCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 907
QY 896 AAGGAGTGGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 955
Db 908 AGGTAGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 967
QY 956 CTCAACATGATACCTGAGAGGAGCAAAATGAGGAGTTCAAGATACAGACCTCATGA 1015
Db 968 CTCAACATGATACCTGAGAGGAGCAAAATGAGGAGTTCAAGATACAGACCTCATGA 1027
QY 1016 AGTGGCTCGGCGTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1075
Db 1028 GGTGGCTCGGCGTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1087

```

OY		1076	CCGGCACTGCGTACTACTGTATGACAAATAATATATAGCTAAGTCATTGGTAACGCTA	11353
OY		1088	CGGTGCCCTCCGGTACTACTATCAGCAAAAATCATATGCACAAGGTGACGGGAAGCGCTA	1147
OY		1136	TGCGTTACAAATTTATTTCCACAGGAATGCGTCAGAGCCCTCCAGCCTCACCTCCAGATTC	1199
OY		1148	TGCGTTACAAAGTTTGACTTCCACGGGATTTGCCAGGCCCTGCAAGCCCCATCCCGGAGTC	1207
Db		1196	ATCCATGTACAAATACCATCAGACCTTCCTACATAGTTCCTAACATGCACACCCCCA	1255
Db		1208	GTCCTCGTAGCAAGTACCCTCCGACCTCCGTACATGGGCTCTATCATCACACCCACCCCCA	1267
OY		1256	GAAATAGAACTTTTAGTGTCCCATCCCCTGCTTGGCCCTTAACCTATCAGAGTTTTT	1315
Db		1268	GAAATAGAACTTTTAGTGTCCCATCCCCTGCTTGGCCCTTAACCTATCAGAGTTTTT	1327
OY		1316	TGCTGCCCTATCATCTACTGGAATTCACAACTGAGAAGCAATACCCAATACAGAGCT	1375
Db		1328	TGCTACCCCGCAACCATACTGTGAATTTGCCGACACTGGGGGACTCTTACCGGAACACTAGGCT	1387
OY		1376	GCCAGCTGCTCATATGCCCTTCCCATCTTGGACACTACTACTA	1417
Db		1388	CCGACGACGACATATGCCCTCTCAGCTGGGACACTACTACTA	1429
RESULT 6				
HOMERG2				
LOCUS		HUMERG2	3166 bp ss-RNA	linear PRI 08-NOV-1994
DEFINITION		Human erg2 gene encoding erg2 protein, complete cds.		
VERSION		M17254		
KEYWORDS		M17254.1 GI:182186		
SOURCE		erg 2 protein.		
ORGANISM		Human cell line COLO 320, cDNA to mRNA, clone lambda 12.		
		Homo sapiens.		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE		1 (bases 1 to 3166)		
AUTHORS		Rao,V.N., Papas,T.S. and Reddy,E.S.		
TITLE		erg, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation, and translation		
JOURNAL		Science 237 (4815), 635-639 (1987)		
MEDLINE		87263429		
PUBMED		3299708		
FEATURES				
source		Location/Qualifiers		
		1..3166		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/map="21q22.3"		
		257..1645		
		/gene="ERG"		
		257..1645		
		/gene="ERG"		
		/note="erg 2 protein"		
		/codon_start=1		
		/protein_id="AA52398.1"		
		/db_xref="GI:182187"		
		/db_xref="GDB:G00-119-884"		
		/translation="MIOIVPPPAHIKEALSVSEDSOLFECAYGTPPIIAKTEMTASS		
		SDDQGSKMSRPYPQQDMLSQPARPTIKCEHPISOYNGSRNSPDECSVAKGSKMG		
		SPDTVGKNISYELEKHMPPEPNNTNRRVTVLPADPTLMSTDHYRQMLEAVKSGYGL		
		DVNLLFQNIIDGKEICKMTKDDEQLRLPSYNADLLSHLHLRETPLPHLLSDVDKRA		
		LONSPLRIHLARNNDIPLYEPPRSAMTHGHETPOSKAOPSSTVPKTEODRPOLDDPY		
		QILDPSTSLRANPGSGOIQLOMFLELLSDSNSCTIMEGTINEFEFTDPDEVARAK		
		KERSKRNNMDYDRALRYLYDKIMTKYNGKRAYAKFDPHGIALOLAPHPRESSIX		
		GESPDLPYMSYNAHLPQKMNFVAHPALPALVTSSFFRAAPRNYSPTGCIYPTNRILP		
		TSHPHSFLGITY"		
		291..298		
		/note="8-bp repeat sequence"		
		509..516		
		/note="8-bp repeat sequence"		
		3143..3149		
		/note="putative"		
		770 c 768 g 713 t		
BASE COUNT		915 a		

Query Match	Best Local Similarity	Matches 1100:	Conservative	0:	Mismatches	245:	Indels	12:	Gaps	2:
63.6%;	Score 921;	DB 9;	Length 3166;							
81.1%;	Pred. No. 2.2e-268;									
227 bp upstream of PvuII site; chromosome 21.										
QY	74	TATTAGAGACATTATTCAGTGTGAGTGAAGACCAAGCTCTTGTGTTGAGTGTGCTACGG	133							
Db	289	TATCAAGAACCCCTTATCAGTTGTGAGTGAAGACCAAGCTCTTGTGTTGAGTGTGCTACGG	348							
QY	134	ATCGCCCCACCTTGGCAAGACAGCAAAATACAGCCCTCTTCACAGTAAATTTGGGCAAC	193							
Db	349	AACGCCACACCTGGCTTAAGACAGAGATACCCGGCTCTCTCCACAGCATTTGACAGAC	408							
QY	194	ATCAAAAGATGAGCCCGCGCTTCCCGACAGAGACTGTTATCAGACCCCGCCAGAGT	253							
Db	409	TTCCAAAGATGAGCCCGCGCTTCCCGACAGAGACTGTTATCAGACCCCGCCAGAGT	468							
QY	254	TACCATTTAAGTGGAGTGAACCCAAACAGTTAATGGGTGCMAGGAATTCACCTGATGA	313							
Db	469	CACCATCAAAATGGATGATTAACCTTAGCAGAGTGAATGGGTGCMAGGAATTCACCTGATGA	528							
QY	314	CTGCAGCGCTGGCAAAAGAGGAAATGGTTAGCGTTCAACACAAATGTTGGGATTAAC	373							
Db	529	ATGCAGTGTGCGCCAAAGCGGGAGAGTGTGGCGACGCCACAGACCGTTGGGATTAAC	588							
QY	374	TGGAAGCTTACATGGAAGAGAAAGCAATATTCGCGCTCCAAATATGACAAACCAATGAACG	433							
Db	589	CGGCAAGCTACATGGAAGAGAAAGCAATATTCGCGCTCCAAATATGACAAACCAATGAACG	648							
QY	434	AGTATTTTGGCAGCAGATCTCTACGTTATGGAGCAGACAGCAATGTACGGCATGTGCTGA	493							
Db	649	AGTATCTGTCGAGAGATCTCTACGTTATGGAGTACAGACCAATGTGCGCATGTGCTGA	708							
QY	494	GTGGGAGGTGAAGAGTATGTGCTTCCAGAGTGGACATCTTGTTGTTCCAGAAATATGA	553							
Db	709	GTGGGAGGTGAAGAGTATGTGCTTCCAGAGTGGACATCTTGTTGTTCCAGAAATATGA	768							
QY	554	TGGGAAGAGTGTGTATAAATGACCAAAAGATGACTTCCAGAGACTCACGCCAGCTATAA	613							
Db	769	TGGGAAGAGTGTGTATAAATGACCAAAAGATGACTTCCAGAGACTCACGCCAGCTATAA	828							
QY	614	CGCAGATATCTCTCTGTACACCTACCTACCTCAGAGAGA-----GAGGAGCCAC	664							
Db	829	CGCGAGACATCTCTCTGTACACCTACCTACCTCAGAGAGACTCTCTTCCACATTTGAC	888							
QY	665	TTTATATTTTCCAAATATCATAGTTTTCGCCAAGAGCAACGCAAGAAATATCAACAGGCC	724							
Db	889	TTTATATATTTTCCAAATATCATAGTTTTCGCCAAGAGCAACGCAAGAAATATCAACAGGCC	948							
QY	725	AGATTTCACCTATAGCAGCAAGCGAGAGATCAGCTGACAGATCACACAGCAATCCCA---C	781							
Db	949	AGATTTCACCTATAGCAGCAAGCGAGAGATCAGCTGACAGATCACACAGCAATCCCA---C	1008							
QY	782	TCAGTCAAAAAGCTACCCACCAACCATCTTCAAGAGTGGCCCAAAACAGAAAGCACAGCGTCC	841							
Db	1009	CCAGTCAAAAAGCTACCTTCAACCATCTTCCACAGTGGCCCAAAACAGAAAGCACAGCGTCC	1068							
QY	842	TCAGTTCAGTCTTATCAGATTTCTTGGACGACCAAGACGCGCTGTGGAATTCAGAGGAG	901							
Db	1069	TCAGTTCAGTCTTATCAGATTTCTTGGACGACCAAGTATGCGCGCTGTGGAATTCAGAGGAG	1128							
QY	902	TGGGACAGATACAGTATGAGGAGTTCTATCTGAGAGTTCTGCGGACAGCTCCCACTCCAA	961							
Db	1129	TGGGACAGATACAGTATGAGGAGTTCTATCTGAGAGTTCTGCGGACAGCTCCCACTCCAG	1188							
QY	962	CTGCATCACCTGGGAGGAGCACAAATGGGGAATTCAGATGACAGACCTGTGTAAGTGGC	1021							
Db	1189	CTGCATCACCTGGGAGGAGCACCAAGGGGAATTCAGATGACAGCTCCGACGAGTGGC	1248							
QY	1022	TCGCGCTTTGGGAGAGAGAAAAGCAACCTTACATGATCTATATACAACTCAGCGCTGC	1081							
Db	1249	CCGCGCTTTGGGAGAGAGAAAAGCAACCTTACATGATCTATATACAACTCAGCGCTGC	1308							



[illegible]

RESULT	8
AB073078	
LOCUS	
DEFINITION	AB073078 2131 bp mRNA linear ROD 17-OCT-2001
ACCESSION	Mus musculus Erg mRNA, mouse homolog of Human ets-related gene Erg,
VERSION	transcript variant 1, complete cds.
KEYWORDS	AB073078.1 GI:16191716
SOURCE	
ORGANISM	Mus musculus cDNA to mRNA, clone:Erg-1.
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 Ozawa,R., Noguchi,H., Taylor,T.D., Takeda,T., Hattori,M. and Sakaki,Y. Mus musculus Erg mRNA Published Only in Database (2001) 2 (bases 1 to 2131) Ozawa,R., Noguchi,H., Taylor,T.D., Takeda,T., Hattori,M. and Sakaki,Y. Direct Submission Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattorigsc.riken.go.jp, URL:http://bgp.gsc.riken.go.jp/, Tel:-81-45-503-9111, Fax:-81-45-503-9170)
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
FEATURES	
Source	Location/Qualifiers 1..2131 /organism="Mus musculus" /db_xref="taxon:10090" /chromosome="16" /clone="Erg-1" 1..2131 /gene="Erg" 148..1536 /gene="Erg" /note="homolog of human ets-related gene (ERG), transcript variant 1" /codon_start=1 /protein_id="BAB69948.1" /db_xref="GI:16191717" /translat="WQIVDPAAHKEALSVSEDSLEFCAYGTPPHLAKTEMTASS SSDYGVKSKSPVPODWLSQAAPARTYIKMECPNSOVNRSRNSPECSYNKGKMWG
gene	
CDS	

Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps	2;
60.9%; Score 881.6; DB 10; Length 2131;	79.3%; Pred. No. 2,1e-256;	0;	0;	269;	12;	2;	
Query	74	TATTAAAGAACCAATTATCACTGTTGAGTGAAGACCAAGCTCTTGTGATGTCCTACGG	133				
Db	180	TATTAAAGAGAGCCTTGTGACGTTGTGAGACGAGAGCACTATTTAGTAGTGCTTCAGG	239				
Qy	134	ATGCCCCCACCCTTGCAGAAAGACAGAAATGACAGCCTCTCTTCAGTGAATATGGGCAAC	193				
Db	240	AACCCACACCTGGCTTAAGACAGAGATGACCGCATCTTCACAGTACATATGGCCAGAC	299				
Qy	194	ATCAAAGATGAGCCCCGCGCTTCCCAAGACAGATCTGTTATCACAGCCCCCGGACAGT	253				
Db	300	ATCCAAAGTATGCTCCACAGATGCTCTTCACAGAGATCTGCTGTCTTAAGCCCCAGCCAGGT	359				
Qy	254	TACCATTAAGATGAGTGTAAACCCAAACAGAGTTAATGGGTCAAGGAATTCACCTGATGA	313				
Db	360	CACCATCAAGATGAGTGAACCCATGACAGGTGAATGGTTCCAGAAATCACCCTGATGA	419				
Qy	314	CTGCAGCGTGGCAAAAGAGGAGAAATGGTTAGCAGTTTCAGACAAATGTTGGGATGAACCTA	373				
Db	420	GTCGAGTGTGAACAAAGTGGGAGATGTTGGGAGCCCGGATCTGTGGGGAATGACCTA	479				
Qy	374	TGGAAGCTACATGGAAGAGAAAGATATTCGCCCTCCCAATATGACAAACAATGAAGCAAG	433				
Db	480	CGGCACTACATGGAAGAGAGAAAGATGTCGCGCTCCCAATATGACAAACAATGAAGCGCAG	539				
Qy	434	AGTATTGTGCGACAGATCCTACGTTATGAGACACAGACCATTAGCGAGTGGCTGA	493				
Db	540	AGTATGATGCTCTGACAGATCTCTACTGTGTGAGACACAGACCATTGCTCGACAGTGGCTGA	599				
Qy	494	GTCGGCAGTGAAGAGATATGCTCTTCCACAGCGTGGACATCTTGTGTTCCAGAACTTGA	553				
Db	600	GTCGGCGGTGAAGAAATATGCGCTCTCCGATGTGAGAGTCTTACTATTATTAAGATATCGA	659				
Qy	554	TGGGAAGAGTGTGTAAATGACCAAAAGATGATCTTCAGAGACTCACGCCGACTATTA	613				
Db	660	TGGGAAGAGTGTGTGAAGATGACAAAGAGATGATCTTCAGAGGCTCGCGGACTATTA	719				
Qy	614	CGCAGATATCTCTGTCACACCTACACCTACCTCAGAGAGA-----GAGGAGCGAC	664				
Db	720	TGCGGACATCTTCTTCTCACATCTCCACTACCTTAGAGAGACTCCCTTCCACATCTGAC	779				
Qy	665	TTTTTATTTTCCAAATACATCAGTTTACCCAGAGAACGCAAGCAAGAAATTAACAAAGCC	724				
Db	780	TTTCGAGTAGAGTTATTAAGGCTTTACAAAACTCTCCACGGTTATATGATGACAGAAACAC	839				
Qy	725	AGATTATACCTTATGAGACAGAGAGAGATACGCTGGACGAGTACAGCCATC---CCAC	781				
Db	840	AGATTATACCTTATGAGCTCTCCACAGAGATCAGCTGGACGGCCACAGCCACCTCACCCC	899				
Qy	782	TCATTCAAAAGCTACCCACCATCATCTTTCACAGCTGCCCAAAACACAAACACAGCGCTC	841				
Db	900	TCAGTCCAAAAGCTGCTAGCCATCTCCCTTTCGACGTGCCAAAACCTAAGACACAGCTCC	959				
Qy	842	TCAGTTAGATCTTATCAGATTTCTTGGACCGACAGCAGCGCTTTCGCAATTCACAGGAG	901				
Db	960	TCAGTTAGATCTTATCAGATTCGCGGACCGACAGACAGCGCGCTTCTTAATCCAGGTAG	1019				
Qy	902	TGGCAGATACAGCTATGACAGTTCTCTACTGAGACTTCTGTGGACAGCTTCAACTGCCA	961				
Db	1020	TGGCAGATACAGCTGTTGGCAGTTCCTGCTGCAATCTCTTCAGACAGCTTCAACTGCCA	1079				







```

Oy 830 AGACGACGTCCTCAGTATAGTCTTATCAGATTCTTGAGCCGACGACGCGCTTGC 889
    |||||||
Db 1020 AGACGACGTCCTCAGTATAGTCTTATCAGATTCTTGAGCCGACGACGCGCTTGC 1079
Oy 890 AATCCAGGAGTGGGAGATACAGTATGAGCTTCTCTAGCTTCTGTCGAGAC 949
    |||||||
Db 1080 TAATCCAGTATGGGAGATACAGTATGAGCTTCTCTAGCTTCTGTCGAGAC 1139
Oy 950 CTCACATCCATGATGATCAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1009
    |||||||
Db 1140 CTCACATCCATGATGATCAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1199
Oy 1010 TGAAGAGTGGCTGGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1069
    |||||||
Db 1200 GGAAGAGTGGCTGGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1259
Oy 1070 ACTGAGCGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1129
    |||||||
Db 1260 GCTGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1319
Oy 1130 ACCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1189
    |||||||
Db 1320 GCGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
Oy 1190 AGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1249
    |||||||
Db 1380 TGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1439
Oy 1250 CCCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1309
    |||||||
Db 1440 CCCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1499
Oy 1310 CTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1369
    |||||||
Db 1500 TTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1559
Oy 1370 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1417
    |||||||
Db 1560 TAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1607

```

```

RESULT 10
AB073079 2133 bp mRNA linear ROD 17-OCT-2001
LOCUS AB073079
DEFINITION Mus musculus Ery mRNA, mouse homolog of Human ets-related gene ERG,
transcript variant 2, complete cds.
ACCESSION AB073079.1 GI:16197542
VERSION 1
KEYWORDS Mus musculus cdna to mRNA, clone:ERG-2.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Oawa,R., Noguchi,H., Taylor,T.D., Takeda,T., Hattori,M. and
Sakaki,Y.
TITLE Mus musculus Ery mRNA
JOURNAL Published only in Database (2001)
REFERENCE 2 (bases 1 to 2133)
AUTHORS Oawa,R., Noguchi,H., Taylor,T.D., Takeda,T., Hattori,M. and
Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Sphiro-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gs.c.riken.go.jp, URL:http://ngp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
FEATURES
SOURCE location/Qualifiers
1..2133
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"

```

```

BASE COUNT 587 a 579 c 531 g 436 t
ORIGIN
Query Match 57.2%; Score 828; DB 10; Length 2133;
Best Local Similarity 76.8%; Pred. No. 4.2e-240;
Matches 1046; Conservative 0; Mismatches 295; Indels 21; Gaps 2;
Oy 74 TATTAGGAACATTATCAGTGTGATGATGAGACCACTCTGTTTATGATGCTTACCG 133
    |||||||
Db 181 TATTAGGAGGCTTGTGACGTTGTGAGCGGAGGACACTATATTAGTGTGCTTACCG 240
Oy 134 ATGCCCCACCTTGCAAGAGAGAAATGACAGCCTCTCTTCCAGTAAATGAGGCAAC 193
    |||||||
Db 241 AACGCCACACCTGCTAAGACAGAGATGACCGCATCTCTTCCAGTAAATGAGGCAAC 300
Oy 194 ATCAAGATGAGCCGCGCTTCCAGCAGAGACTGTTATACAGGCCCGCGGACAGT 253
    |||||||
Db 301 ATCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Oy 254 TACCATTAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 313
    |||||||
Db 361 CACCATTAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
Oy 314 CTGAGCGTGGCAAAAGAGGAGAAATGCTTACGATGATGATGATGATGATGATGATG 373
    |||||||
Db 421 GTGCACTGTGTAACAAAGGTGGAGATGATGATGATGATGATGATGATGATGATG 480
Oy 374 TGAAGCTACATGGAAGAGAGATATTCGCTCCAAATATGACCAACATGAAAGCAG 433
    |||||||
Db 481 CGGAGCTACATGGAAGAGAGATATTCGCTCCAAATATGACCAACATGAAAGCAG 540
Oy 434 AGTTATTTGCTCCAGCAGATCTTACGTTATGAGAGCAGACCATGATGAGTGTGGA 493
    |||||||
Db 541 AGTATGCTCCGCAATATCTTCTGATGAGAGCAGACCATGATGAGTGTGGA 600
Oy 494 GTGAGCAGTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 553
    |||||||
Db 601 GTGAGCAGTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Oy 554 TGGGAAAGATGTTGTAAATGACCAAGATGATGATGATGATGATGATGATGATGAT 613
    |||||||
Db 661 TGGGAAAGATGTTGTAAATGACCAAGATGATGATGATGATGATGATGATGATGAT 720
Oy 614 CGCAGATATCTCTCTGACACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 673
    |||||||
Db 721 TGGCAGATATCTCTCTGACACCTTACCTTACCTTACCTTACCTTACCTTACCT 780
Oy 674 TCCAAATATCATGATTTACCAAGAGCA-----CGCAAGATATAC 715
    |||||||
Db 781 TTCCGATGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Oy 716 AACAGGCCAGATTTATCATGAGCAGGAGATGATGATGATGATGATGATGATGATG 775
    |||||||
Db 841 AGGGGGGCGAGCTTTATTTTCCAAATATCTTACCTTACCTTACCTTACCTTACCT 900

```



QY	655	-GAGAGCACCCTTTTATTTTTTCCAAATAAATACGTTTAAACCGAGAGACGCAAGAT	712
Db	1019	AGGAGGTGCCTCTTTTATTTTTTCCAAATCATTCCGTTTATCAAGACGCAAAACAAAGAT	1078
QY	713	AACAACAAGCCGAGATTTTACCTTATGAGCAAGCGAGAGATCAGCGTAGAGAGTCACAG	772
Db	1079	CCCCAGCAGGCAAGATTTATCATATGAGGCCCTCAAGGAGATCTGCTTGACAAATCAT--	1136
QY	773	CCATCCCACTCAGTCAAAAGCTAACCAACATCATCTTCAACAGTGGCCCAAAACAGAAG	832
Db	1137	-CTGACACACCATCTCAAAAGCATCTCAAC---CATCTACCAAGTTCCTCCAAACAGAGA	1192
QY	833	CCAGGCTCCTCAGTTAGATCCCTTATCAATCTTGAGACGACGACGACGCTTGCAAA	892
Db	1193	CCCCAGGCGACAGCTAGATGATCTCTTATCAATCTTGAGCAACAGCAGCAGCTTGCAAA	1255
QY	893	TCCAGGAGTGGGCGAGATACAGCTATGAGCAGTTCCTACTGAGAGCTTCTGTGCGACGTC	952
Db	1253	TCCAGGAGTGGGCGAGATTCAGCTCTGTGCAATTCCTTACTGGAAGTCTTGGAATGCTC	1312
QY	953	CAACTCCCACTGCTCACCCTGGGAGGCGACAATATGGGAGATTCAAGATGACAGACCTGA	1012
Db	1313	CAACTCCCACTGCTCATTACTTGGGAGGAGCAACATGGAATTTAAGATGACGATCCCGA	1372
QY	1013	TGAAGTGCCTGGCGCTGGGAGAGAGAGAAAGCAACCTTAACATGACCTATGACCAACT	1072
Db	1373	TGMAATGGCAGACAGCTTGGGGGAGAGAGAAAGCAACCAACATGACATGACACACT	1432
QY	1073	CAGCCGTCACCTTGCTCTACTATGACAAAAATATTATGACTAAAGTTCATGTTAAAG	1132
Db	1433	CAGCCGTCACCTTGCTCTACTACTAGATAAAATATTATGACTAAAGTTCATGTTAAAG	1492
QY	1133	CTATGCCCTACAAATTTGATTTCCAGGGAATGGCTCAGGCCCTCCAGCTCAGCTCCAGA	1192
Db	1493	CTATGCCCTACAAATTTGATTTCCAGGGAATGGCTCAGGCCCTCCAGCTCAGCTCCAGA	1552
QY	1193	ATCATTCACATGACAAATACCATATGACAGCCTCCCTACATGAGTTCCTACATGACACAC	1252
Db	1553	ATCATTCACATGACAAATACCATATGACAGCCTCCCTACATGAGTTCCTACATGACACAC	1612
QY	1253	CCAGAGATGAACTTTGAGCTCCCATCCCTCTGCTTTGGCCGTAACCTCATCCAGCTT	1312
Db	1613	ACAGAGATGAACTTTGAGCTCCCATCCCTCTGCTTTGGCCGTAACCTCATCCAGCTT	1672
QY	1313	TTTGGCTGCGCCCTATCATCTACTGGAATTCACCAAGTGGAGGCACTCTACCCCATACAG	1372
Db	1673	CTTGGCAACCCCTATGATGATCTGGAATTCACCAAGTGGAGTATTTATCCAAATACCTG	1732
QY	1373	GCTGCGAGCTGCTATATGATGCTTCCCATCTTTGGACACCTACCTACCTAAGGGGGA	1425
Db	1733	GCTGCGAGCTGCTATATGATGCTTCCCATCTTTGGACACCTACCTATTTAAATCAAGA	1785
LOCUS	HUMERG11	3126 bp	mRNA
DEFINITION	Human erg protein (ets-related gene)	mRNA, complete cds.	
ACCESSION	M21535	M1390	
VERSION	M21535.1	GI:182182	
KEYWORDS	erg protein.		
SEGMENT	1 of 2		
SOURCE	Human, cell line COLO 320, cDNA to mRNA, Lambda-7.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Reddy, E.S., Rao, V.N. and Paps, T.S.		
TITLE	The erg gene: a human gene related to the ets oncogene		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 84 (17), 6131-6135 (1987)		
MEETING	87317608		
PubMed	3476934		
FEATURES	Location/Qualifiers		

	source	1..3126	/organism="Homo sapiens"
		/db_xref="taxon:9606"	
CDS		195..1286	
		/note="ergl protein"	
		/codon_start=1	
		/protein_id="AAA3811.1"	
		/db_xref="GI:182185"	
		/translation="MVSGSDPYGMNYSYMEKCHMPNNMTNERRIVLPADPLTMSI DHVROMLEMAKEYEELPYNVILIFPNIDGKELCRTKDREFQRLPTSYADIILSLHLH LREITPHLISDDVDKAUQLNSPRMLHANPDPIEPFRSRMGTGHHTPSKAPQAQ PSTVTKTEDORQDLDPYLIGPTSSRLANPSQIQLOWFLEILSDSNSSCTIMEP TNGEKMRDPEVARMERGEKSKPMNYMDKLISRALRYDYDNIMTKYGVAKAFEDT HGIALQHPHEPESLSLYRKSPDIPMGYSVYHHQMFNVADHPALPVTSSEFFAAPV PYMNSPGTGYTPNERLPLSHMPSHLGTY"	
BASE COUNT	928 a	732 c	725 g
ORIGIN	1 bp upstream from EcoRI site.	741 t	
Query Match	52.5%; Score 760.2; DB 9;	Length 3126;	
Best Local Similarity	80.7%; Pred. No. 1.8e-219;		
Matches 917; Conservative	0; Mismatches 208; Indels 12; Gaps 2;		
Dn	294 TCACGAATTCACCTCGATGACTGCAGCCTGGCAAAAGAGGAAAAATGTTAAGCATCTCA	353	
Dn	150 TCAAGAACCTCCCTCGATGAAATGACGTGTGGCAAAGCGGGAATAATGTGGGACGCCA	209	
Dn	354 GACAATGTGGGATTAACCTATGGAAGATCTACATGGAAGAAGACATATTCCGCTCCAAT	413	
Dn	210 GACACCGTTGGGATTAACCTACAGCGCATACATGAGAGGAAGACATGCCACCOCACAAAC	269	
Dn	414 ATGACACCAATGAACGAAGAGTTATTGTCCAGCACAGATCCTACGTTATGAGACACAGAC	473	
Dn	270 ATGACACGAAGACGACCGCAGAGTTATCTGCGCAGCAGATCCTACGCTATGAGATACAGAC	329	
Dn	474 CATGTACGGCAGTGCTGAGTAGGGCAGTGAAGAAGATGTGGTCTCCAGAGTGGACATC	533	
Dn	330 CATGTGGCGCACTGGCTGGAGTGGCGGTGAAAGAAATATGGCTTCCAGAGCTCAACATC	389	
Dn	534 TTGTGTTCCAGAACATGATGGGAAAGAGTTGTGTTAAATGACCAAAAGATGACTCCAG	593	
Dn	390 TTGTTATTCAGAAATCGATGATGGGGAAGACATGTGCAGATGACCAAGACAGACTCCAG	449	
Dn	594 AGACTCAGCGCGAGCTATACGCGAGATTTCTCTGTCACACACTACACTACTCAGAGAG	653	
Dn	450 AGGCTACCCCCAGGTACACGCGCAGATCTCTCTCATCATCTCCACTACTCCTCAGAGAG	509	
Dn	654 A-----GAGGAGCCACTTTTATTTTTTCCAAATACATGATGATTACCGAAGACAGC	704	
Dn	510 ACTCCTCTTCCACATTTGACTTCAGATATGTGTATTAAGCTTACAAACATCTCCACAG	569	
Dn	705 CAAAGATTAACCAACAGGCCAGATTTACCTTATGACGAAGCGAGAGATCAGCGTGCAGC	764	
Dn	570 TTTAATGACGTGTAGAAACACAGATTTTACATATGAGCCCCCAGAGATCAGCGTGCAGC	629	
Dn	765 AGTCACACCACATCCCA--CTCAGTCAAAAGCTTACCCCAACATCATCTTCAACAGTGGCC	821	
Dn	630 GTTACGCGCCACCCACGCCGCCAGTGTGAAAGCTGTCTAACCATCTCTTCCACAGTGGCC	689	
Dn	822 AAAACAGAAAGACACAGCGTCTCAGTTAGATCCTTATCAGATTCTTGAACGACGACGAC	881	
Dn	690 AAAATGTAAGACACAGCGTCTCAGTTAGATCCTTATCAGATTCTTGAACCAAGATGAC	749	
Dn	882 CGTCTTGAATCCAGGAGTGGGACAGATTAAGCTATGAGGAGTTCTTACTGGAGACTTCTG	941	
Dn	750 CGCCTTGAATTCACAGGACGAGTGGCAGATTCAGGCTTGGCAGTTCTCTGAGACTCTCTG	809	
Dn	942 TCGGACACCTTCAACTCCAACTGCATCCTCTGGAGGAGGACAAATATGGGAGATTCAAGATG	1001	
Dn	810 TCGGACAGCTCCAACTCCAGCTGCATCCTCTGGGAAAGCACCAACGGGAGTTCAAGATG	869	
Dn	1002 ACAGACCTGATGAAGTGGCTCGGCGTTGGGGAGAGAGAAAGCAAACTTAACATGAAC	1061	
Dn	870 ACAGATCCCGACGAGTGGCCCGCGCTGGGGAGAGGAGAGCAAAACCCCAACATGAAC	929	

Oy	1062	TATGACAACAGCCCGGCGACCTTGGCTACGATGACAAAAAATATGACAAAGTT	1121
Db	930	TACGATTAAGCTCAGCCCGGCCCTCCGTTACTACTATGACAGAAACATCATCCAGAGTC	989
Oy	1122	CATGCTAAACGCTATGGCTTACAAATTTGATTTCCACGGGAATCGCTCAGGCCCTCAGGCT	1181
Db	990	CATGGGAAGGCTACGCTTACAAAGTTGACATTCCACGGGATCGGCCAGGCCCTCGAGGCC	1049
Oy	1182	CACCTTCGAGATTCATCATGTACAAATACCATCAGACCTCCCTTACATAGTTCTTAC	1241
Db	1050	CACCCCGGGAGTATCTGTGTACAGATCCCTCAGACCTCCGTCATGAGTGGCTCCAT	1109
Oy	1242	CATGCACACCCCGGAAAGATGACATCTTTGTACCTCCCATCCCGCTGGCTTGGCGTAAC	1301
Db	1110	CAGCGCCACCCACAGAAAGATTAACCTTTGTGGCGCCACCCCTCCAGCCCTCCCGTGACA	1159
Oy	1302	TCATTCAGCTTTTGTGTCGCCCTAATCCATACACTGGAATTCACCAACATGAGGACATCAC	1361
Db	1170	TCCTTCAGATTTTTTTGGTGGCCCAACCCATACGATGAAATTCACCAAGTGGGGATATATAC	1229
Oy	1362	CCCAATACACAGCTGCCAGCTGCTCATATGCTTCCATCTTTGGCAGCTACTACTAA	1418
Db	1230	CCCAACACTAGGCTCCACCCAGCAGCAATATGCTTCTCATCTGCGGACACTTACTACTAA	1286
RESULT 13			
LOCUS	DRE249590	1436 bp	linear VRT 03-MAR-2000
DEFINITION	Danio rerio mRNA for fil-1 protein (fil-1 gene).		
ACCESSION	AJ249590		
VERSION	AJ249590.1	GI:6006476	
KEYWORDS	fil-1 gene; FIL-1 protein.		
SOURCE	Danio rerio.		
ORGANISM	Danio rerio.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
AUTHORS	1 (bases 1 to 1436)		
TITLE	Brown, L.A., Rodaway, A.R., Schilling, T.F., Jowett, T., Ingham, P.W., Patient, R.K. and Sharrocks, A.D.		
JOURNAL	Insights into early vasculogenesis revealed by expression of the Ets-domain transcription factor Fil-1 in wild-type and mutant zebrafish embryos		
MEDLINE	Mech. Dev. 90 (2), 237-252 (2000)		
REFERENCE	20108585		
PUBMED	10640707		
AUTHORS	2 (bases 1 to 1436)		
TITLE	Sharrocks, A.D.		
JOURNAL	Direct Submission		
FEATURES	Submitted (29-SEP-1999) Sharrocks A.D., School of Biological Sciences, University of Manchester, Oxford Road, Manchester, M13, 9PT, UNITED KINGDOM		
SOURCE	Location/Qualifiers		
gene	1. 1436		
CDS	/organism="Danio rerio"		
	/db_xref="taxon:7955"		
	/dev_strage="20-24hr embryo"		
	45. 1400		
	/gene="fil-1"		
	45. 1400		
	/gene="fil-1"		
	/function="transcription factor"		
	/codon_start=1		
	/evidence=experimental		
	/product="Fil-1 protein"		
	/protein_id="CAB56832.1"		
	/db_xref="GI:6006477"		
	/translation="MDGTIKELSVSEDSLEFPYAAAPLTKDTMASGTQDYGO		
	TKHNPPIPOEWINOPVRNVKREYHINGSRSPDVGSKGNKMWGGREASOMAY		
	TGWDEKCAAPPNMTNTERVRYIPADPSLMSPDYRWRLDMAIKETGLQETDPMFHS		
	TDGELEKMSKDFLRLLTSVNTFVLSHLYNKRESSSTSYNTPSHADDSPLAARD		
	DASDVAARRGWSNNMHSIGKSPVYSQVSKNDPDPDPDPIQIIGPSSRLANG		

Query Match	Best Local Similarity	39.5%;	Score 571.2;	DB 5;	Length 1436;	
Matches	901;	Conservative	0;	Mismatches 443;	Indels 24;	Gaps 4;
QY	63	ATGGCAAGCACTATTAAAGCAAGCATTTATCAAGTGTGTAAGTAAGACCAAGTCTTGTTTGAAG	122			
Db	45	ATGACAGCAATTTAAAGAGAGCGCCTGTCACTGGTAAGTAAGAACCAATCTCTGTGTGACG	104			
QY	123	TGTGGCTTAAGCATGGCCCCACCTTGCA---AAGACAGAAATGACACGCTCTCTTCCAGT	179			
Db	105	CTCTCATACGCGCTGCTGCTCTTATACCAAGACAGCATGACTGCTCCGGGACCTCAG	164			
QY	180	GAATATGGGCAAAACATCAAAAGATAGAGCCCGCGCGTTCGCCAGCAGAGTGTATTACAG	239			
Db	165	GATTACGGCCACAGACACAAATATCAACCCCATTTCCCTTCAGCAAGATGGATTATATCAG	224			
QY	240	CCCCCGGCGCAGATTACCATTTAAGTGGAGTGTAAACCAACCAAGGTTAATGGGTCAAG	299			
Db	225	CC-----GGTCCGGGTAAAGGTCAAGGAGAGTATGACACATCAACGAGATCCAGA	275			
QY	300	AATTCACCTGATGACTGCGAGGTGGCAAAAGAGAGAAATGTTAGCATTACAGACAT	359			
Db	276	GAGTGCCTGGGTGACTCGACGCGTGGGAAATATGCAATATAATGTGTGGCGGACCGAACGC	335			
QY	360	GTTGGAGTGAACATATGAAGCTACATGTGAGAGAAAGCAATATTCGCTCCAAATATGACA	419			
Db	336	TCTCAGATGAACCTATCTGCTACATGATGATGAATATGCTCGCCCTCCAAACATATGACG	395			
QY	420	ACCAATAGAGACAGATTATGTGGCCACACATCCTACGTTATGAGGACAGACCATATTA	479			
Db	396	ACCAATAGAGAGAGATGATGATCTCCCGCAGACCCGCTCTGTGTGTCTCCGGATACGTG	455			
QY	480	CGGAGTGTGGTGGAGTGGGAGTGAAGAGATATGGTCTTCCAGACGTGACATCTTGTTG	539			
Db	456	CGGAGTGTGTAGACTGGGCCATTAAAGAGATACGGTCTTCAGAGATGACACAGGCGATG	515			
QY	540	TTTCCAGAACATTTGATGGGAAGAGTGTGTAAATAGACCAAAAGATGATTTCCACAGACTC	599			
Db	516	TTTTCACAGCACAGATGGAAGAGACTCTGCAAGATGAGACAAAGACAGATTTCTTAGACTTC	575			
QY	600	ACGCGGAGCTTATACGACGATATCTCTCTGTCACACACTACACTACTCTAGAGAGAGAGA	659			
Db	576	ACCAAGCGTTTATAACCGAGAGTCTCTCTCTCACATCTCAATTTACTAGGGAAGAATAG-	634			
QY	660	GCCACTTTTATTTTCCAAATATACATCACTTATACCCAGAGACAGCAAGAATTAACACA	719			
Db	635	--CTCATCTGAATATATACACACGCCATCTACGCTGACGATCCCTCCGCTGGCTGCA	692			
QY	720	AGGCGAGTTTACCTTTATGAGCAGCGAGGAGATACAGGTGGAGAGATCACAGCATCC	779			
Db	693	AAACACGATGCTGTTATGATGCTGTACGAGAGAGAGGAGCTGTGTGAACACATGCACAGC	752			
QY	780	ACTCAGTCAAAAGGTACCCCAACACATCATCTTCAACAGTGCCTCAAAACAGAGACCAAGCT	839			
Db	753	GGAAAGAGCTTCCCAACAGTGTGCTCTCAGAGCTGTCTCAAAACCCGAGATACAGCTGT	812			
QY	840	CCTCAGTTAAGTCTTATACATATCTTGGACACGACACGACCGCTTGTGCAAAATTCACAGG	899			
Db	813	CCTCAGCGAGATCCTTATACATATTTAGAGACCCACACAGCACTGTGATTCACCAAGC	872			
QY	900	AGTGGCAGATACAGCATATGCAATGCTTCTACTGAGACTTCTGTGGAGACAGTCCAAATCC	959			
Db	873	TCAGTGTAGATCCACAGCTGTGGCAATTTCTGCTGGAGCTCTGTCCGAGAGGCCAACGCC	932			
QY	960	AACGTGATCACCTTGGAGGCGCAAAATGGGAGTTTCAAGATACAGACCTTGATGAAGTG	1019			

Db	Accession	Source	Organism	Reference Authors	Journal Title	Journal
Db	993	GACGTGCATACCTGGGAAGGAACCAAGGAGTTCAGATGATGACGACCCAGACGAGTGG	1079			
Qy	1020	GCTCGGCGCTTGGGGAGAGAGAAAACAAACCTAACTGAATGAACTGACAACTCAGCCGT	1079			
Db	993	GCACGGCCCTGGGGGGAACGCAAGAGCAAAACCCAACTGAATGAACTGACAACTGAGCCGA	1052			
Qy	1080	GCACCTTGGCTACTATGACAAAAATTTATGACTAAAGTTTCATGTTAAAGCTATGCC	1139			
Db	1053	GCCCTGCGCTACTACTAGACAGAAACATCATGACCAAAAGTGCAGAGGACGCTGACGCC	1112			
Qy	1140	TACAATTTGATTTCCACAGGAATGCGCTGAGGCCCTCCAGCCCTGACCCCTCCAGAAATCATCC	1199			
Db	1113	TACAAGTTGCACTTCCAGCGCATGCGCATGCGCCAGCGCCCTGCAACCCCACTCCAGATGCCACC	1172			
Qy	1200	ATGTACAAATACCCATCAGACCTCCCTCCATGAGTTCCATCCATGACGACCCCGAGAG	1255			
Db	1173	ATGTACAAATACCCAGGACGATGCGCGTACGCTTACCTATGACGACCCAGACGAGAG	1232			
Qy	1260	ATGAACTTTGAGCTCCCATCCCTCTCTTGGCCCGTAACCTATCCAGCTTTTCT	1319			
Db	1233	GTCACCTTCGCTCTCCACCTCCGCAATGCGCCGACCTCCCTCCATCTTCCTCGT	1292			
Qy	1320	GCCTTATTCATCTGGAATTCACCACTGAGGACATCTACCCCAATACGAGGCTGCCA	1379			
Db	1293	CCCAACACCCATACCTGAGAGCTCCCGACAGAGATTTTACCCCAATCCAGAGCTGCCA	1352			
Qy	1380	-----GCTGCTCATATGCTCTCCCTCCATCTTGGACCTACTACTAA	1418			
Db	1353	CGCCATGCCAATTCACAGCTGCTTACACCTTGATGTACTACTAA	1400			
RESULT 14						
MEFLI	MEFLI	1729 bp	RNA	linear	ROD 02-SEP-1991	
LOCUS	Mouse Fl1-1	mouse	Fl1-1	retroviral	integration site.	
DEFINITION	X59421					
ACCESSION	X59421.1	GI:50974				
KEYWORDS	c-ets-1 gene; ets gene family; fl1-1 gene; integration site.					
SOURCE	Mus musculus.					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE	1 (bases 1 to 1729)					
AUTHORS	Letwin, K.					
TITLE	Direct Submission					
JOURNAL	Submitted (08-MAY-1991) K. Letwin, Samuel Lunenfeld Res Inst at Mount Sinai Hospital, Div of Mol and Developmental Biol, 600 University Avenue, Toronto Ontario M5G 1X5, CANADA					
REFERENCE	2 (bases 1 to 1729)					
AUTHORS	Ben-David, Y., Giddens, E.B., Letwin, K. and Bernstein, A.					
TITLE	Erythroleukemia induction by Friend murine leukemia virus: Insertional activation of a new member of the ets gene family, Fl1-1, closely linked to c-ets-1					
JOURNAL	Genes Dev. 5 (6), 908-918 (1991)					
MEDLINE	91257578					
PUBMED	2044959					
FEATURES	Location/Qualifiers					
source	1..1729					
	/organism="Mus musculus"					
	/strain="balb/c"					
	/db_xref="taxon:10090"					
	/cell_line="CB7"					
	/cell_type="Friend erythroleukemia"					
	/germline					
	1..1729					
	/gene="Fl1-1"					
	1..1729					
	/gene="Fl1-1"					
	/evidence=experimental					
	245..1603					
	/gene="Fl1-1"					
	/codon_start=1					
	/product="retrovirus integration site"					

BASE COUNT	466 a	484 c	436 g	343 t	
ORIGIN					
Query Match	39.4%	Score 570.2;	DB 10;	Length 1729;	
Best Local Similarity	66.8%	Pred. No. 1.1e-161;			
Matches 917;	Conservative 0;	Mismatches 423;	Indels 33;	Gaps 6;	
QY	63	ATGGCAAGCACATTTAATGAAGAACATTTACATGCTGTGATGAAGACCAAGTCTTGTGTGAG	122		
DB	245	ATGAGACGGAGCATTAAGAGAGGCGTCTGCTGTGTGATGACATCAGTCCCTTTTGAT	304		
QY	123	TGTGCTTACGATG---CCCCACCTTGCAACACACAAATGACAGCCCTCTTCCAGT	179		
DB	305	TCACCATACGAGGAGGGGAGCCCAATCTCCCAAGGCAATATGACTGCTTGGGGAGTCT	364		
QY	180	GAATATGGGCAAAATCAATCAAGATATAGCCCGGCGCTTCCGACGAGACTGGTTATCCAG	239		
DB	365	GACTACGGGACGCCCCACAAATATACCCCTGCCACCGCACGGAGTGATACACAG	424		
QY	240	CCCCGGGACAGTTTACCATTAAGATGGAGTGTAAACCCAAACCAAGTTAATGGGTCAAG	299		
DB	425	CC-----AGTAGAGTCAATGTCAAGGGGAGTATGACACATGAATGGATCCAGG	475		
QY	300	AATTACCTGATGACTGACGCGTGCGCAAAAGGAGGAAATGTTTACAGTTCAGACAAT	359		
DB	476	GAGTCTCCGGTGAGCTGACGTGCAGCAAAATGTAAACAAGTGGTGGGCGGAGGCGAAGCC	535		
QY	360	GTTGGGATGAATCTATGAGAGCTATACATGGAAGAGAGA---TATTCCGCGCTCCAAATATG	416		
DB	536	AACCCCATGAACATTAATAGCTATCATGATGATGAGAGAAGCGGCCCTCTCTCCCAACATG	595		
QY	417	ACAACCAATGACAAAGATTAATGTTGCCACGACAGATCTTACGTTATGAGACACAGACAT	476		
DB	596	ACCAACCAAGAGGAGAGATATTGTGCTGCTGCGACAGACCCACACCTGTGGACACAGAGAC	655		
QY	477	GTAAGGCAAGTGTGAGTGGGCGGACGTAGAGAGAGTATGTTCTTCCAGACGTGCACATCTTG	536		
DB	656	GTTGCAACAGTGGCTGAGTGGGCTATTAAGAAATATACGATTAATGATGATGACACTTCC	715		
QY	537	TGTGTTCCAGAACATTTGATGGGAAGATTTGTGTAATATGACCAAAAGATGACTTCCAGAGA	596		
DB	716	TTCTTCCAGAACATGATGATGGAAGAAATTTGTAAATATGAACAAGGAGGACTTCTCCGA	775		
QY	597	CTCACGCGGAGCTTTAAGCAGGATATCTCTCTGACACACTACACTACCTCTCAGAGAGAGA	656		
DB	776	GCCACCTCCGCTCAACACAGAGTCTGTGTGCGACCTCAGTTACTCTCAGAGGAAG-	834		
QY	657	GGAGGCACTTTATTTTCCAAATACATCAATCAAGTTTACCAGAGCAACGCAAAAGATTAACA	716		
DB	835	-----TTACATGCTGGGCTTAAACACACCTCCCATACAGACAGTCTCAGACATGAAT	889		
QY	717	ACAAGGCCAATTTACTTTATGAGACGACGAGAGATACACCGTGGACGAGTCA---GC	773		
DB	890	GTCAGAGGAAGACCTTTATGATCTGTGAGAGAGAGACATGGAACAATTAATATGAAC	949		
QY	774	CATCCCACTAGTCAAAAGCTACCCACACATCTTCAACAGTGGCCCAAAAGAGAAG	833		
DB	950	TCTGGCCTCAACAAAGTCTCTCTCTTGGAGAGATACAGACCATGGGCAAGAACACTGAG	1009		
QY	834	CAGGCTCCTAGTTAGTCCCTTATCAGATTTCTTGAGACGACGACGACGCTCTTGAAT	893		

Db 1010 CACGGGCCCCAGAGATCCCTTATCAAGATCTGGGGCCCAACAGACACCCCTAGCAAC 1069  
 QY 894 CCAAGGAGTGGGAGATACAGATATGGCACTTCTTCTGAGGAGTCTGGGAGAGCTTC 953  
 Db 1070 CTTGGAGTGGGAGATACAGATCTGGCACTTCTTCTGAGGAGTCTGGGAGAGCTTC 1129  
 QY 954 AACTCCAACTGCATCACCTGGGAGGAGCAAAATGGGAGTTCAGATGACAGACCTGAT 1013  
 Db 1130 AACGCCAGTGTATACCTGGGAGGAGGAGGAGGAGTTCAGAAATGAGGAGACCTGAT 1189  
 QY 1014 GAAGTGGCTGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1073  
 Db 1190 GAGGTGGCCAGGGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1249  
 QY 1074 AGCGGTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133  
 Db 1250 AGCGGGGCTCCGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309  
 QY 1134 TATGCTTACAAATTTGATTTTCAAGGAGTGCCTGAGGCTCCAGCCTCCAGCTCCAGAA 1193  
 Db 1310 TATGCTTACAAATTTGATTTTCAAGGAGTGCCTGAGGCTCCAGCCTCCAGCTCCAGAA 1369  
 QY 1194 TCAATCCTTACAAATTTGATTTTCAAGGAGTGCCTGAGGCTCCAGCCTCCAGCTCCAGAA 1253  
 Db 1370 ACATCCATGATCAAGATATCCCTGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1429  
 QY 1254 CAGAAGATGAATTTGATTTTCAAGGAGTGCCTGAGGCTCCAGCCTCCAGCTCCAGCTT 1313  
 Db 1430 CAGAAGTGAATTTGATTTTCAAGGAGTGCCTGAGGCTCCAGCCTCCAGCTCCAGCTT 1489  
 QY 1314 TTTGCTGCTTATCCATTTGATTTTCAAGGAGTGCCTGAGGCTCCAGCCTCCAGCTT 1373  
 Db 1490 TTTGAGGAGTATCAAGATATCCCTGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1549  
 QY 1374 CTGCGAGTGTCTATA-----TGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1417  
 Db 1550 GTCCCGCCGCTATCAACCCACAGCTGCTTCACTAGGAGACTACTACTA 1602

RESULT 15  
 A36461 2916 bp DNA linear PAT 05-MAR-1997  
 LOCUS A36461  
 DEFINITION Sequence 2 from Patent WO9323549.  
 ACCESSION A36461.1 GI:2293779  
 VERSION A36461.1  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2916)  
 Aurias,A., Delattre,O., Desmaza,C., Melot,T., Peter,M.,  
 Plouhastel,B., Thomas,G. and Zucman,J.  
 NUCLEIC ACID CORRESPONDING TO A GENE OF CHROMOSOME 22 INVOLVED IN  
 RECURRENT CHROMOSOMAL TRANSLOCATIONS ASSOCIATED WITH THE  
 DEVELOPMENT OF CANCEROUS TUMORS  
 Patent: WO 9323549-A 2 25-NOV-1993;  
 CENTRE NAT RECH SCIENT (FR)  
 Other publication FR 2691475 931126  
 Other publication JP 8500964T 960206.  
 COMMENT  
 FEATURES  
 source  
 location/Qualifiers  
 1..2916  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /dev\_stage="ADULT"

BASE COUNT 844 a 686 c 663 g 723 t  
 ORIGIN  
 Query Match 39.0%; Score 563.8; DB 6; Length 2916;  
 Best Local Similarity 66.5%; Pred. No. 1e-159;  
 Matches 913; Conservative 0; Mismatches 427; Indels 33; Gaps 6;

QY 63 ATGGCAAGCATATTAAGAGCATTAATGAGTGTGAGTGAAGACCAAGCTCTTTTGGAG 122  
 Db 121 ATGAGAGGAGCTATTAAGAGAGCTGTGCTGGTGTGAGCAGACAGCTCTCTTTGAC 180  
 QY 123 TGTGCTAGGAGATG---CCCACTTTGCAAAAGCAGAAATGACAGCTCTCTTCCAGT 179  
 Db 181 TCAGCGTACGAGACGGGAGCCATCTCCCAAGGCGCAGCTGCTGGGAGTCTT 240  
 QY 180 GAATATGAGCAAAATCAATCAAGATGAGCCCGGCTTCCAGAGAGCTGTTATACAG 239  
 Db 241 GACTAGGAGGAGCCCAAGATCAACCCCTCCACACAGAGAGTGTATCAATCAG 300  
 QY 240 CCCCCGCGCAGGTTACATTAAGATGAGTGTAAACCAAGGTTAATGGTCAAGG 299  
 Db 301 CC-----AGTAGAGGTCAACGTCAGCGGAGTATGACATGATGATTCAGG 351  
 QY 300 AATTCACTATATGCTTACACCGTGGCAAAAGAGGAGAAATGGTTAGCAATTCAGCAAT 359  
 Db 352 GAGTCTCCGGTGAAGCTGACGCTTACGCAAAATGAGCAAGCTGGGAGGAGGCTC 411  
 QY 360 GTTGGGATGAATATGGAAGCTACATGAGAGAGAGCAT---ATTCCGCTCCAAATATG 416  
 Db 412 AACCCATGATACATCAAGAGCTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 471  
 QY 417 ACAACCAATGAAGAGAGTATATGTCAGCAGATCTTACGTTATGAGAGCAGACCAT 476  
 Db 472 ACCACCAAGAGAGAGAGTATGTCAGCAGATCTTACGTTATGAGAGAGAGAGAGAT 531  
 QY 477 GTACGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 536  
 Db 532 GTAGGCAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 591  
 QY 537 TTTGCTCAGAACTTATGAGGAGAGTGTGTAATGACCAAAATGATCTCCAGAGA 596  
 Db 592 TTTTTCAGAACTATGATGTGCAAGAGAGTGTGTAATGACCAAAATGAGAGAGAGAG 651  
 QY 597 CTCAGCGCAGATATTAAGCAGATATCTCTCTGACACCTTACACTACAGAGAGA 656  
 Db 652 GGCACCAAGCTCTTACCAAGAGAGTGTGTCACACTCTGTTACAGTACAGAGAGAG- 710  
 QY 657 GGAGCCACTTTATTTTCCAAATATCATGTTTACCCAGAGAGCAGCAAGCAATTAACA 716  
 Db 711 -----TTTACGCTGGCTTATTAATACACTCCCAACAGCAATCTTCACATTTGAGT 765  
 QY 717 ACAAGGCAATTTTACCTTATGAGCAAGAGAGATCAAGCTGAGAGTCAACACCAT 776  
 Db 766 GTCAAGAGAGACCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 825  
 QY 777 CC---CACTCAGTCAAAAGTATACCAACATCTTTCACAGTGTCCCAAAACAGAGAC 833  
 Db 826 TCTGGCTCTCAAAAGTCTCTCCCTTGGAGGGGACCAAAAGTCAAGTAAATTAAGATGAG 885  
 QY 834 CAGCGTCTAGTTAGTCTTATGATTTCTGAGACCGAGCAGAGCGCTTGGCAAT 893  
 Db 886 CAAGGCCCCAGCAATCTGATCAAGATCTGAGGCCCCGAGCAGAGTGGCTTAGCAAC 945  
 QY 894 CCAAGGAGTGGCAGATTAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 953  
 Db 946 CCGTGAAGGAGGAGATCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1005  
 QY 954 AACTCCAACTGCATCACCTGGGAGGAGCAGCAAAATGGGAGTTCAGATGACAGACCTGAT 1013  
 Db 1006 AACGCCAGTGTATACCTGGGAGGAGGAGCAGCAAGGAGTTCAGAAATGAGAGCCCGAT 1065  
 QY 1014 GAAGTGGCTGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1073  
 Db 1066 GAGGTGGCCAGGGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1125  
 QY 1074 AGCGGTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133  
 Db 1126 AGCGGGGCTCCGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185  
 QY 1134 TATGCTTACAAATTTGATTTTCAAGGAGTGCCTGAGGCTCCAGCTCCAGAA 1193



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 07:39:19 ; Search time 32.1776 seconds

(Without alignments)  
9948.861 Million cell updates/sec

Title: US-09-902-772-1\_COPY\_650\_660

Perfect score: 11

Sequence: 1 agagagagagag 11

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb.pa:\*

2: gb.htg:\*

3: gb.in:\*

4: gb.om:\*

5: gb.ov:\*

6: gb.pat:\*

7: gb.ph:\*

8: gb.pl:\*

9: gb.pr:\*

10: gb.ro:\*

11: gb.sts:\*

12: gb.sy:\*

13: gb.un:\*

14: gb.vi:\*

15: em.ba:\*

16: em.fun:\*

17: em.hum:\*

18: em.in:\*

19: em.mu:\*

20: em.om:\*

21: em.or:\*

22: em.ov:\*

23: em.pat:\*

24: em.ph:\*

25: em.pl:\*

26: em.ro:\*

27: em.sts:\*

28: em.un:\*

29: em.vi:\*

30: em.htg.hum:\*

31: em.htg.in:\*

32: em.htg.other:\*

33: em.htg.mus:\*

34: em.htg.pln:\*

35: em.htg.pod:\*

36: em.htg.man:\*

37: em.htg.vtl:\*

38: em.sy:\*

39: em.htg.hum:\*

40: em.htg.mus:\*

41: em.htg.other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	20	AX326907	AX326907 Sequence
2	11	100.0	21	AX244546	AX244546 Sequence
3	11	100.0	21	AX404317	AX404317 Sequence
4	11	100.0	21	AX404318	AX404318 Sequence
5	11	100.0	22	DOGP34701	DOGP34701 Sequence
6	11	100.0	22	AB6933	AB6933 Sequence
7	11	100.0	22	AR048345	AR048345 Sequence
8	11	100.0	22	AR079236	AR079236 Sequence
9	11	100.0	23	AR118353	AR118353 Sequence
10	11	100.0	24	MM884	MM884 Sequence
11	11	100.0	25	AX115095	AX115095 Sequence
12	11	100.0	32	AR060668	AR060668 Sequence
13	11	100.0	32	AR102813	AR102813 Sequence
14	11	100.0	36	AR084537	AR084537 Sequence
15	11	100.0	36	AX113835	AX113835 Sequence
16	11	100.0	42	A06938	A06938 F. domesticu
17	11	100.0	43	AR011922	AR011922 Sequence
18	11	100.0	43	AR017793	AR017793 Sequence
19	11	100.0	43	AR077198	AR077198 Sequence
20	11	100.0	43	E38119	E38119 DNA elongat
21	11	100.0	51	ECY17989	ECY17989 Enchelyopus
22	11	100.0	63	AR088300	AR088300 Sequence
23	11	100.0	63	AR140197	AR140197 Sequence
24	11	100.0	63	AR198558	AR198558 Sequence
25	11	100.0	63	BD008437	BD008437 Fibroblas
26	11	100.0	63	MPY17987	MPY17987 Microtest
27	11	100.0	89	MVB2LF	MVB2LF M.vision B-2
28	11	100.0	90	AR037143	AR037143 Sequence
29	11	100.0	90	AX35642	AX35642 Sequence
30	11	100.0	90	165754	165754 Sequence
31	11	100.0	99	MUSCERBI	MUSCERBI Mus musculu
32	11	100.0	104	AR037140	AR037140 Sequence
33	11	100.0	104	165751	165751 Sequence
34	11	100.0	105	MVB2LB	MVB2LB M.vision B-2
35	11	100.0	105	9	9 HUMH2LAH
36	11	100.0	108	AY009753	AY009753 Human micro
37	11	100.0	111	6	6 AR020453
38	11	100.0	111	6	6 AR020454
39	11	100.0	111	6	6 AR18349
40	11	100.0	111	6	6 AR118350
41	11	100.0	111	6	6 AR122315
42	11	100.0	111	6	6 AR122316
43	11	100.0	111	6	6 I80021
44	11	100.0	111	6	6 I80022
45	11	100.0	112	4	4 AF276820

## ALIGNMENTS

RESULT 1

AX326907

LOCUS AX326907

DEFINITION Sequence 103 from Patent WO0178894.

ACCESSION AX326907

VERSION AX326907.1 GI:18097618

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

synthetic construct.

artificial sequences.

REFERENCE

1 Keith, T.

AUTHORS

TITLE Novel human gene relating to respiratory diseases, obesity, and inflammatory bowel disease

JOURNAL Patent: WO 0178894-A 103<25-OCT-2001>

Pred. No. is the number of results predicted by chance to have a

Genome Therapeutics Corp. (US)  
 Location/Qualifiers  
 1. .20  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="Primer"

BASE COUNT 6 a 3 c 9 g 2 t

Query Match 100.0%; Score 11; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4e+04;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
 |||||  
 2 AGAGAGAGAG 12

RESULT 2  
 AX244546/c 21 bp DNA linear PAT 28-SEP-2001  
 LOCUS Sequence 12 from Patent WO0166793.  
 DEFINITION AX244546  
 ACCESSION AX244546  
 VERSION AX244546.1 GI:15859473  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Vitis vinifera.  
 Vitis vinifera  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Vitaceae; Vitis  
 1 (bases 1 to 21)  
 baiges Blanco, I.M., Schaeffner, A.R. and Mas, A.  
 Nucleic acid and oligonucleotides derived therefrom for  
 specifically amplifying and specifically detecting aquaporin genes  
 from Vitis spec  
 Patent: WO 0166793-A 12-13-SEP-2003;  
 GSF-Forschungszentrum fuer Umwelt und Gesundheit GmbH (DE)  
 Location/Qualifiers  
 1. .21  
 /organism="Vitis vinifera"  
 /db\_xref="taxon:29760"

BASE COUNT 2 a 9 c 1 g 9 t

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+04;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
 |||||  
 20 AGAGAGAGAG 10

RESULT 3  
 AX404317 21 bp DNA linear PAT 14-JUN-2002  
 LOCUS Sequence 143 from Patent WO0224747.  
 DEFINITION AX404317  
 ACCESSION AX404317  
 VERSION AX404317.1 GI:21437598  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 synthetic construct.  
 synthetic construct  
 artificial sequences.  
 1  
 Brinkmann, U. and Hoffmeyer, S.  
 Polymorphisms in human genes of cardiovascular regulators and their  
 use in diagnostic and therapeutic applications  
 Patent: WO 0224747-A 143-28-MAR-2002;  
 Epidauros Biotechnologie AG (DE)  
 Location/Qualifiers  
 1. .21  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"

FEATURES  
 source

/note="artificial sequence"  
 BASE COUNT 7 a 4 c 10 g 0 t

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+04;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
 |||||  
 11 AGAGAGAGAG 21

RESULT 4  
 AX404318 21 bp DNA linear PAT 14-JUN-2002  
 LOCUS Sequence 144 from Patent WO0224747.  
 DEFINITION AX404318  
 ACCESSION AX404318  
 VERSION AX404318.1 GI:21437599  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 synthetic construct.  
 synthetic construct  
 artificial sequences.  
 1  
 Brinkmann, U. and Hoffmeyer, S.  
 Polymorphisms in human genes of cardiovascular regulators and their  
 use in diagnostic and therapeutic applications  
 Patent: WO 0224747-A 144 28-MAR-2002;  
 Epidauros Biotechnologie AG (DE)  
 Location/Qualifiers  
 1. .21  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="artificial sequence"

BASE COUNT 0 a 10 c 4 g 7 t

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+04;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
 |||||  
 11 AGAGAGAGAG 1

RESULT 5  
 DOGP34701/c 22 bp DNA linear MAM 28-FEB-1996  
 LOCUS Dog (Clone: CXX-347) primer for STS 347, 5' end.  
 DEFINITION L24232  
 ACCESSION L24232  
 VERSION L24232.1 GI:401891  
 KEYWORDS  
 PCR identification; PCR primer; STS.  
 1 of 2  
 Canis familiaris (library: E. Ostrander, in pBluescript+) adult  
 spleen DNA.  
 Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 1 (bases 1 to 22)  
 Ostrander, E.A., Maza, F.A., Yee, M. and Rine, J.  
 One hundred and one new simple sequence repeat-based markers for  
 the canine genome  
 Mamm. Genome 6 (3), 192-195 (1995)  
 95268214  
 MEDLINE 7749226  
 PUBMED  
 COMMENT  
 Submitted by:  
 Fred Hutchinson Cancer Research Center  
 Transplantation Biology Dept  
 1124 Columbia; Mailstop M318  
 Seattle, WA 98104, USA  
 e-mail: EAOstrander@hl.gov  
 PCR Buffer: PCR buffer (Perkin-Elmer/Cetus)

PCR Profile: Denaturation: 94 degrees C for 1.00 minute  
Annealing: 55 or 59 degrees C for 0.45 minutes  
Polymerization: 74 degrees C for 1.00 minutes  
PCR Cycles: 33  
Final Extension: 74 degrees C for 5.00 minutes.

FEATURES  
source  
1. .22  
/organism="Canis familiaris"  
/db\_xref="taxon:9615"  
/tissue\_type="spleen"  
/dev\_stage="adult"  
/tissue\_lib="E. Ostrander, in pbluescript+"

BASE COUNT 0 a 10 c 3 g 9 t  
ORIGIN  
primer\_bind 1. .22

Query Match 100.0%; Score 11; DB 4; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.9e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
DB 19 AGAGAGAGGAG 9

RESULT 6  
LOCUS A86933 22 bp DNA linear PAT 22-JAN-2000  
DEFINITION Sequence 24 from Patent WO9838306.  
ACCESSION A86933  
VERSION A86933.1 GI:6735717  
KEYWORDS

SOURCE unidentified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 22)  
AUTHORS Dolganov, G.  
TITLE TRANSCRIPTS ENCODING IMMUNOMODULATORY POLYPEPTIDES  
JOURNAL PATENT: WO 9838306-A 24 03-SEP-1998;  
GENELABS TECH INC (US)  
FEATURES  
source  
1. .22  
Location/Qualifiers

BASE COUNT 2 a 12 c 0 g 8 t  
ORIGIN  
Query Match 100.0%; Score 11; DB 6; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.9e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
DB 15 AGAGAGAGGAG 5

RESULT 7  
LOCUS AR048345 22 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 24 from patent US 5821091.  
ACCESSION AR048345  
VERSION AR048345.1 GI:5970688  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 22)  
AUTHORS Dolganov, G.  
TITLE Method of identifying activated T-cells  
JOURNAL Patent: US 5821091-A 24 13-OCT-1998;  
FEATURES  
source  
1. .22  
Location/Qualifiers

BASE COUNT 2 a 12 c 0 g 8 t  
ORIGIN  
Query Match 100.0%; Score 11; DB 6; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.9e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
DB 15 AGAGAGAGGAG 5

RESULT 8  
LOCUS AR079236 22 bp DNA linear PAT 31-AUG-2000  
DEFINITION Sequence 24 from patent US 5965427.  
ACCESSION AR079236  
VERSION AR079236.1 GI:10005982  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 22)  
AUTHORS Dolganov, G. and Novikov, A.  
TITLE Human RAD50 gene and methods of use thereof  
JOURNAL Patent: US 5965427-A 24 12-OCT-1999;  
FEATURES  
source  
1. .22  
/organism="unknown"

BASE COUNT 2 a 12 c 0 g 8 t  
ORIGIN  
Query Match 100.0%; Score 11; DB 6; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.9e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
DB 15 AGAGAGAGGAG 5

RESULT 9  
LOCUS AR118353 23 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 5 from patent US 6140493.  
ACCESSION AR118353  
VERSION AR118353.1 GI:14099259  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 23)  
AUTHORS Dower, W.J., Barrett, R.W., Gallop, M.A. and Needels, M.C.  
TITLE Method of synthesizing diverse collections of oligomers  
JOURNAL Patent: US 6140493-A 5 31-OCT-2000;  
FEATURES  
source  
1. .23  
/organism="unknown"

BASE COUNT 11 a 0 c 12 g 0 t  
ORIGIN  
Query Match 100.0%; Score 11; DB 6; Length 23;  
Best Local Similarity 100.0%; Pred. No. 3.9e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11  
DB 10 AGAGAGAGGAG 20

RESULT 10  
MMB84

LOCUS MMB84 24 bp DNA linear ROD 07-MAR-1997  
DEFINITION M.musculus immunoglobulin heavy chain CDR3 region (germlinal centre  
B8 DNA #4).  
ACCESSION X67390  
VERSION X67390.1 GI:50118  
KEYWORDS diversity region; immunoglobulin; immunoglobulin heavy chain;  
joining region; variable region.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 24)  
AUTHORS Jacob, J.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUL-1992) J. Jacob, University of Maryland School of  
Medicine, Dept of Microbiology & Immunology, 655 W Baltimore St,  
Baltimore MD 21 201, USA  
REFERENCE 2 (bases 1 to 24)  
AUTHORS Jacob, J. and Kelsoe, G.  
TITLE In situ studies of the primary immune response to  
(4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for  
peritoneal lymphoid sheath-associated foci and germinal centers  
J. Exp. Med. 176 (3), 679-687 (1992)  
JOURNAL 92381435  
MEDLINE 1512536  
REFERENCE 3 (bases 1 to 24)  
AUTHORS Jacob, J., Przybela, J., Miller, C. and Kelsoe, G.  
TITLE In situ studies of the primary immune response to  
(4-hydroxy-3-nitrophenyl)acetyl. III. The kinetics of V region  
mutation and selection in germinal center B cells  
J. Exp. Med. 178 (4), 1293-1307 (1993)  
JOURNAL 93389394  
MEDLINE 8376635  
COMMENT See also X67341-7, X67349-9L, J00522, J00529-30, J00532-37 &  
J00539.  
FEATURES  
source location/Qualifiers  
1. 24  
/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/cell\_line="germlinal centre B8"  
/cell\_type="B cells"  
/tissue\_type="spleen"  
/dev\_stage="day 10"  
<1..>24  
/note="CDR3"  
/codon\_start=1  
/protein\_id="CAA47802.1"  
/db\_xref="GI:1333873"  
/translation="ARERSSYL"  
CDS  
BASE COUNT 8 a 3 c 8 g 5 t  
ORIGIN  
Query Match 100.0%; Score 11; DB 10; Length 24;  
Best Local Similarity 100.0%; Pred. No. 3.8e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGGAG 11  
|||||  
Db 4 AGAGAGAGGAG 14  
RESULT 11  
AX115095/c 25 bp DNA linear PAT 11-MAY-2001  
LOCUS AX115095  
DEFINITION Sequence 218 from Patent WO0129262.  
ACCESSION AX115095  
VERSION AX115095.1 GI:14032037  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 25)

AUTHORS Picoult-Newburg, L. and Pohl, M.  
TITLE Genotyping reagents, kits and methods of use thereof  
JOURNAL Patent: WO 0129262-A 218 26-APR-2001;  
Orchid Biosciences, Inc. (US)  
FEATURES  
source location/Qualifiers  
1. 25  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Primer"  
BASE COUNT 4 a 12 c 0 g 9 t  
ORIGIN  
Query Match 100.0%; Score 11; DB 6; Length 25;  
Best Local Similarity 100.0%; Pred. No. 3.8e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGGAG 11  
|||||  
Db 22 AGAGAGAGGAG 12  
RESULT 12  
AR060668/c 32 bp DNA linear PAT 29-SEP-1999  
LOCUS AR060668  
DEFINITION Sequence 42 from patent US 5840839.  
ACCESSION AR060668  
VERSION AR060668.1 GI:5987118  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 32)  
AUTHORS Wang, R.-F. and Rosenberg, S.A.  
TITLE Alternative open reading frame DNA of a normal gene and a novel  
human cancer antigen encoded therein  
Patent: US 5840839-A 42 24-NOV-1998;  
JOURNAL Location/Qualifiers  
1. 32  
/organism="unknown"  
FEATURES  
source  
BASE COUNT 7 a 9 c 7 g 9 t  
ORIGIN  
Query Match 100.0%; Score 11; DB 6; Length 32;  
Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAGAGGAG 11  
|||||  
Db 29 AGAGAGAGGAG 19  
RESULT 13  
ARI02813/c 32 bp DNA linear PAT 14-FEB-2001  
LOCUS ARI02813  
DEFINITION Sequence 42 from patent US 6087110.  
ACCESSION ARI02813  
VERSION ARI02813.1 GI:12814401  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 32)  
AUTHORS Wang, R.-F. and Rosenberg, S.A.  
TITLE Alternative open reading frame DNA of a normal gene and a novel  
human cancer antigen encoded therein  
Patent: US 6087110-A 42 11-JUL-2000;  
JOURNAL Location/Qualifiers  
1. 32  
/organism="unknown"  
FEATURES  
source  
BASE COUNT 7 a 9 c 7 g 9 t  
ORIGIN  
Query Match 100.0%; Score 11; DB 6; Length 32;

Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
|||||  
Db 29 AGAGAGAGAG 19

## RESULT 14

AR084537 36 bp DNA linear PAT 01-SEP-2000  
LOCUS AR084537  
DEFINITION Sequence 26 from patent US 5981185.  
ACCESSION AR084537  
VERSION AR084537.1 GI:10011308  
KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 36)  
AUTHORS Matsun,R.S., Coassin,P.J., Rampal,J.B. and Caskey,C.Thomas.  
TITLE Oligonucleotide repeat arrays  
JOURNAL Patent: US 5981185-A 26-09-NOV-1999;  
FEATURES Location/Qualifiers  
source 1..36  
/organism="unknown"

BASE COUNT 18 a 0 c 18 g 0 t  
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 36;  
Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
|||||  
Db 18 AGAGAGAGAG 28

## RESULT 15

AX113835/c 36 bp DNA linear PAT 01-MAY-2001  
LOCUS AX113835  
DEFINITION Sequence 11 from Patent WO0127295.  
ACCESSION AX113835  
VERSION AX113835.1 GI:13940018  
KEYWORDS

SOURCE synthetic construct.  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 36)  
AUTHORS Schaedendorf,D., Paschen,A., Chakraborty,T. and Dommann,E.  
TITLE Recombinant attenuated listerias for immunotherapy  
JOURNAL Patent: WO 0127295-A 11-19-APR-2001;  
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts  
(DE)

FEATURES Location/Qualifiers  
source 1..36  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Primer"

BASE COUNT 8 a 5 g 10 t  
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 36;  
Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11  
|||||  
Db 35 AGAGAGAGAG 25

Search completed: July 28, 2003, 11:20:41  
Job time : 37.1776 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 07:41:54 ; Search time 429.733 Seconds

(Without alignments)  
7582.941 Million cell updates/sec

Title: US-09-902-772-1

Perfect score: 1447  
Sequence: 1 gaattcgcgaacgaataat.....gaaagaagcgccaaagaaa 1447

Scoring table: IDENTITY\_NNC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N\_Geneseq\_101002.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1447	100.0	1447	20	AA26551
2	1356	93.7	1528	20	AA26552
3	563.8	39.0	2938	14	AA050644
4	562.2	38.9	2957	24	ABR84139
5	544.6	37.6	2954	22	AAH02915
6	387	26.7	567	22	ABA48124
7	387	26.7	567	22	ABA66003
8	387	26.7	567	22	ABA33090
9	387	26.7	567	22	AAK14426

C 10	387	26.7	567	22	AAK40160	Human bone marrow
C 11	387	26.7	567	22	AAI20932	Probe #10865 for g
C 12	387	26.7	567	22	AAI6176	Probe #14862 used t
C 13	387	26.7	567	22	AAI06643	Probe #6634 used t
C 14	387	26.7	567	24	ABSL4215	Human genome-deriv
C 15	354.6	24.5	473	22	ABA50193	Human breast cell
C 16	354.6	24.5	473	22	ABA68128	Human foetal liver
C 17	354.6	24.5	473	22	ABA35152	Probe #13618 for g
C 18	354.6	24.5	473	22	AAK16512	Human brain expres
C 19	354.6	24.5	473	22	AAK42265	Human bone marrow
C 20	354.6	24.5	473	22	AAI48339	Probe #17025 used t
C 21	354.6	24.5	473	22	AAI08696	Probe #8687 used t
C 22	325	22.5	533	21	AAI01342	Human secreted pro
C 23	320.4	22.1	1890	14	AA050662	Human Hum-Pli-1 ge
C 24	276.6	19.1	549	24	ABK44114	CDNA #54 encoding
C 25	255	17.6	420	22	ABA45056	Human breast cell
C 26	255	17.6	420	22	ABA55529	Human foetal liver
C 27	255	17.6	420	22	ABA25239	Probe #3705 for ge
C 28	255	17.6	420	22	AAK03760	Human brain expres
C 29	255	17.6	420	22	AAK29224	Human bone marrow
C 30	255	17.6	420	22	AAI35180	Probe #3866 used t
C 31	255	17.6	420	22	AAI03692	Probe #3683 used t
C 32	255	17.6	454	22	ABA42995	Human breast cell
C 33	255	17.6	454	22	ABA53411	Human foetal liver
C 34	255	17.6	454	22	ABA23187	Probe #1653 for ge
C 35	255	17.6	454	22	AAK01681	Human brain expres
C 36	255	17.6	454	22	AAK27129	Human bone marrow
C 37	255	17.6	454	22	AAI11718	Probe #1651 for ge
C 38	255	17.6	454	22	AAI33025	Probe #1711 used t
C 39	255	17.6	454	22	AAI01646	Probe #1637 used t
C 40	255	17.6	454	22	ABSO1685	Human genome-deriv
C 41	235.6	16.3	318	21	AAK21875	Human secreted pro
C 42	180	12.4	1752	22	AA513672	CDNA encoding Ratt
C 43	180	12.4	1752	24	AAI38753	Rat lambda73 CDNA
C 44	170.6	11.8	899	23	ABLI2097	Drosophila melanog
C 45	167.8	11.6	553	22	ABA61050	Human foetal liver

# ALIGNMENTS

RESULT 1	AA26551	standard; DNA; 1447 BP.
ID	AA26551	
AC	AA26551	
XX	14-JUN-1999	(first entry)
DT		
XX		
DE	DNA encoding chicken C-11 protein.	
XX		
KW	Chicken; C-11 protein; cell calcification inhibiting activity;	
KW	cell calcification inhibiting agent; c-erg protein; arthritis deformans;	
KW	ossification; spinal column ligament; ss.	
XX		
OS	Gallus sp.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	63..1418
FT		/*tag= a
XX		
PN	JP11075871-A.	
XX		
PD	23-MAR-1999.	
XX		
PF	29-MAY-1998;	98JP-0166076.
XX		
PR	20-JUN-1997;	97US-0050294.
PR	18-JUN-1997;	97US-0878177.
XX		
PA	(CHUS ) CHUGAI PHARM CO LTD.	
PA	(UYPE-) UNIV PENNSYLVANIA.	
XX		

DR WPI: 1999-257708/22.  
 DR P-PSDB: AAY01520.  
 XX An active protein for inhibiting cell calcification - useful for  
 PT measuring the calcification of a cell, for diagnosing arthritis  
 PT deformans or ossification of spinal column ligament  
 XX  
 PS Disclosure: Page 7-8; 15pp; Japanese.  
 XX  
 CC The present sequence encodes a chicken C-11 protein which has cell  
 CC calcification inhibiting activity. The specification also describes  
 CC a cell calcification inhibiting agent containing c-ery protein  
 CC (AAY01521). The proteins are used for measuring the calcification of a  
 CC cell, for diagnosing arthritis deformans or ossification of spinal column  
 CC ligament.  
 CC  
 XX Sequence 1447 BP; 440 A; 374 C; 317 G; 316 T; 0 other;  
 Query Match 100.0%; Score 1447; DB 20; Length 1447;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGGAGCAATATTTATTTAGCAATTTATTTAGGATCAATATTTGATGACA 60  
 DB 1 GAATTCGGGAGCAATATTTATTTAGCAATTTATTTAGGATCAATATTTGATGACA 60  
 QY 61 TTATGGCAAGCACTATTATTAAGCAATTTAGTGTGATGAAGACCACTCTGTTTG 120  
 DB 61 TTATGGCAAGCACTATTATTAAGCAATTTAGTGTGATGAAGACCACTCTGTTTG 120  
 QY 121 AGGTGCTTACGATGCGCCCTTGTGCAAGAGCAAGAAATGACAGCTCTCTTCCAGTG 180  
 DB 121 AGGTGCTTACGATGCGCCCTTGTGCAAGAGCAAGAAATGACAGCTCTCTTCCAGTG 180  
 QY 181 AATATGGGCAAAATCAATGAAGAGCGCGCTCCCGACAGAGACTGGTTATACAGC 240  
 DB 181 AATATGGGCAAAATCAATGAAGAGCGCGCTCCCGACAGAGACTGGTTATACAGC 240  
 QY 241 CCCCAGGCAAGTTCACATTAAGATGAGTGTAAACCCAAACAGGTTAATGGGTCAAGA 300  
 DB 241 CCCCAGGCAAGTTCACATTAAGATGAGTGTAAACCCAAACAGGTTAATGGGTCAAGA 300  
 QY 301 ATTACCTGATGACTGCGAGGTGGCAAAAGAGGAAATGTTAGCAGTTTCAGACATG 360  
 DB 301 ATTACCTGATGACTGCGAGGTGGCAAAAGAGGAAATGTTAGCAGTTTCAGACATG 360  
 QY 361 TTGGGATGAACATCTGGAAGCTACATGAGAGAGAAAGCATATTCGCTCCAAATATGCAA 420  
 DB 361 TTGGGATGAACATCTGGAAGCTACATGAGAGAGAAAGCATATTCGCTCCAAATATGCAA 420  
 QY 421 CCAATGAACGAGAGTATTTGTGCGACAGATCTCTACGTTATGAGACACAGACCATGTAC 480  
 DB 421 CCAATGAACGAGAGTATTTGTGCGACAGATCTCTACGTTATGAGACACAGACCATGTAC 480  
 QY 481 GGCAGTGGCTGGAGTGGCGAGTGAAGAGATGCTTCCAGACGTGGAGCATCTTGTGT 540  
 DB 481 GGCAGTGGCTGGAGTGGCGAGTGAAGAGATGCTTCCAGACGTGGAGCATCTTGTGT 540  
 QY 541 TCCAGAAACATGATGAGGAAAGAGTGTGTTAAATGACAAAGATGACTTCACAGACTCA 600  
 DB 541 TCCAGAAACATGATGAGGAAAGAGTGTGTTAAATGACAAAGATGACTTCACAGACTCA 600  
 QY 601 CGCGGAGCTATAACGAGATATCTCTGTCACACCTACACTACCTCTCAGAGAGAGAG 660  
 DB 601 CGCGGAGCTATAACGAGATATCTCTGTCACACCTACACTACCTCTCAGAGAGAGAG 660  
 QY 661 CCACCTTTATTTTCCAAATACATCATCTTACCCAAAGAGCAAGCAAGATTAACACAA 720  
 DB 661 CCACCTTTATTTTCCAAATACATCATCTTACCCAAAGAGCAAGCAAGATTAACACAA 720  
 QY 721 GGCCAGATTTACCTTATGAGCAAGGAGAGATCAGCGTGGACGAGTCAAGCATCCCA 780  
 DB 721 GGCCAGATTTACCTTATGAGCAAGGAGAGATCAGCGTGGACGAGTCAAGCATCCCA 780

QY 781 CTCAGTCAAAAGCTACCCAAACATCATCTTCAACAGTGCCTCAAAACAGAGACGCGTC 840  
 DB 781 CTCAGTCAAAAGCTACCCAAACATCATCTTCAACAGTGCCTCAAAACAGAGACGCGTC 840  
 QY 841 CTCAGTTAGATCTTATCAATCTTGTGACCGGACGACGACCGCTTGTGCAATTCAGGGA 900  
 DB 841 CTCAGTTAGATCTTATCAATCTTGTGACCGGACGACGACCGCTTGTGCAATTCAGGGA 900  
 QY 901 GTGGGCGATATACAGCTATGAGCAGTCTCTACTGAGAGTCTGTGAGACAGCTCCAACTCCA 960  
 DB 901 GTGGGCGATATACAGCTATGAGCAGTCTCTACTGAGAGTCTGTGAGACAGCTCCAACTCCA 960  
 QY 961 ACTGCATCACCTGGGAGGAGCAAAATGGGAGTTCAAGATGACAGACCTGATGAAGTGG 1020  
 DB 961 ACTGCATCACCTGGGAGGAGCAAAATGGGAGTTCAAGATGACAGACCTGATGAAGTGG 1020  
 QY 1021 CTGGGCGTTGGGAG 1080  
 DB 1021 CTGGGCGTTGGGAG 1080  
 QY 1081 CACTGCTACTACTATGACAAAATTTATGACTTAAAGTTCAATGATGAAGCTATGCTT 1140  
 DB 1081 CACTGCTACTACTATGACAAAATTTATGACTTAAAGTTCAATGATGAAGCTATGCTT 1140  
 QY 1141 ACAATTTGATTTCCAGGAAATGCTCAGGCGCTCCAGCCTCCAGAAATCATCCA 1200  
 DB 1141 ACAATTTGATTTCCAGGAAATGCTCAGGCGCTCCAGCCTCCAGAAATCATCCA 1200  
 QY 1201 TGTACAAATACCCATCAGACCTCCCTCAGTATGAGTCTTACATGACACACCCCAAGA 1260  
 DB 1201 TGTACAAATACCCATCAGACCTCCCTCAGTATGAGTCTTACATGACACACCCCAAGA 1260  
 QY 1261 TGAACCTTGTAGTCCCATCCCTCCCTGCTTGGCCGTAACCTATCAGCTTTTGTCTG 1320  
 DB 1261 TGAACCTTGTAGTCCCATCCCTCCCTGCTTGGCCGTAACCTATCAGCTTTTGTCTG 1320  
 QY 1321 CCCCTAATCATCTGGAATTCACCACTGAGGACATCTACCCCAATACAGGCTGCGAG 1380  
 DB 1321 CCCCTAATCATCTGGAATTCACCACTGAGGACATCTACCCCAATACAGGCTGCGAG 1380  
 QY 1381 CTGCTCATATGCTTCCCATCTTGGCAGCTTACTACTATGAGGAGAGAGAGAGAGAG 1440  
 DB 1381 CTGCTCATATGCTTCCCATCTTGGCAGCTTACTACTATGAGGAGAGAGAGAGAGAG 1440  
 QY 1441 AAGAAAA 1447  
 DB 1441 AAGAAAA 1447

RESULT 2  
 AAX26552  
 ID AAX26552 standard; DNA: 1528 BP.  
 AAX26552:  
 14-JUN-1999 (first entry)  
 DNA encoding chicken c-ery protein.  
 Chicken; C-11 protein; cell calcification inhibiting activity;  
 cell calcification inhibiting agent; c-ery protein; arthritis deformans;  
 ossification; spinal column ligament; ss.  
 Gallus sp.  
 Key Location/Qualifiers  
 CDS 63..1499  
 FT /\*tag= a  
 JP11075871-A.  
 23-MAR-1999.



XX 29-MAY-1998; 98JP-0166076.  
XX 20-JUN-1997; 97US-0050297.  
XX 18-JUN-1997; 97US-0878177.  
XX (CHUS-) CHUGAI PHARM CO LTD.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX WPI; 1999-257708/22.  
XX P-PSDB; AAY01521.  
XX An active protein for inhibiting cell calcification - useful for  
XX measuring the calcification of a cell, for diagnosing arthritis  
XX deformans or ossification of spinal column ligament  
XX  
XX Disclosure; Page 8-9; 15pp; Japanese.  
XX  
XX The present sequence encodes a chicken c-ery protein. The specification  
XX also describes a chicken C-11 protein (AAY01520) which has cell  
XX calcification inhibiting activity and a cell calcification inhibiting  
XX agent containing c-ery protein. The proteins are used for measuring the  
XX calcification of a cell, for diagnosing arthritis deformans or  
XX ossification of spinal column ligament.  
XX  
XX Sequence 1528 BP; 464 A; 395 C; 330 G; 339 T; 0 other;  
SO  
Query Match 93.7%; Score 1356; DB 20; Length 1528;  
Best Local Similarity 94.7%; Pred. No. 0;  
Matches 1447; Conservative 0; Mismatches 0; Indels 81; Gaps 1;  
QY 1 GAATTCGGCAAGAAATATTTATTTAGCAATTTATGAGATCAATATCTTGATACA 60  
DB 1 GAATTCGGCAAGAAATATTTATTTAGCAATTTATGAGATCAATATCTTGATACA 60  
QY 61 TTATGGCAACATTTAAGGAAGCATTTATCAGTGTGAGTGAAGACAGCTCTTGTG 120  
DB 61 TTATGGCAACATTTAAGGAAGCATTTATCAGTGTGAGTGAAGACAGCTCTTGTG 120  
QY 121 AGTGTGCTTACGATGCGCCACCTTGTCGCAAGACAGAAATGACAGCTCTTGTG 180  
DB 121 AGTGTGCTTACGATGCGCCACCTTGTCGCAAGACAGAAATGACAGCTCTTGTG 180  
QY 181 AATATGGCAAAATCAAGATGAGCCCGGCGTCCCGCAGAGATGTTATACAGC 240  
DB 181 AATATGGCAAAATCAAGATGAGCCCGGCGTCCCGCAGAGATGTTATACAGC 240  
QY 241 CCCCGGCAAGTTACCATTAAGTGAAGTGAACCCAAACAGGTTAATGGGTCAGGA 300  
DB 241 CCCCGGCAAGTTACCATTAAGTGAAGTGAACCCAAACAGGTTAATGGGTCAGGA 300  
QY 301 ATTACCTGATGCTGAGCGGTGGCAAGAGGAGAAATGTTAGCAGTTCAAGCATG 360  
DB 301 ATTACCTGATGCTGAGCGGTGGCAAGAGGAGAAATGTTAGCAGTTCAAGCATG 360  
QY 361 TTGGGATGAAGTGAAGTGAAGTGAAGAGAGATTTCCGCTCCAAATATGCAAA 420  
DB 361 TTGGGATGAAGTGAAGTGAAGTGAAGAGAGATTTCCGCTCCAAATATGCAAA 420  
QY 421 CCAATGAAGAGATTTATGTCGACAGATCTTACGTTATGAGAGACAGACATTTAC 480  
DB 421 CCAATGAAGAGATTTATGTCGACAGATCTTACGTTATGAGAGACAGACATTTAC 480  
QY 481 GGCAGTGGTGGAGTGGGAGTGAAGAGATGTTTCCAGAGCTGATCTTGTGT 540  
DB 481 GGCAGTGGTGGAGTGGGAGTGAAGAGATGTTTCCAGAGCTGATCTTGTGT 540  
QY 541 TCCAGAACATTTGGAAGAGTGTGTTAAATGACCAAGATGACTTCCAGAGACTCA 600  
DB 541 TCCAGAACATTTGGAAGAGTGTGTTAAATGACCAAGATGACTTCCAGAGACTCA 600  
QY 601 CGCGGAGCTTAAAGCAGATATCTCTCTGACACCTACACTGCTAGAGAGA----- 654  
DB 601 CGCGGAGCTTAAAGCAGATATCTCTCTGACACCTACACTGCTAGAGAGCTCTC 660  
QY 655 ----- 654  
DB 661 TTCACATTTGACTTCAGATGATGTGATAAGGCTTACAAAACCTCTCCAGGTTAATGC 720  
QY 655 -----GAGGAGCCACTTTTATTTTCCAAATATGACTGATTTACCCAGAG 699  
DB 721 ATGCTAGAAACACAGGAGAGCCACTTTTATTTTCCAAATATGACTGATTTACCCAGAG 780  
QY 700 CAACGCAAAAGATTAACAAGAGGCGAGATTACCTTATGAGCAAGCAGAGATCAGCGT 759  
DB 781 CAACGCAAAAGATTAACAAGAGGCGAGATTACCTTATGAGCAAGCAGAGATCAGCGT 840  
QY 760 GAGAGATCACAGGATCCAGCTGATCAAAAGGTACCAACCATCTTCAACAGTGC 819  
DB 841 GAGAGATCACAGGATCCAGCTGATCAAAAGGTACCAACCATCTTCAACAGTGC 900  
QY 820 CCAAAACAGAAAGACACAGCTCTTATGATGATCTTATCAGATTTCTTGAGCCAGCA 879  
DB 901 CCAAAACAGAAAGACACAGCTCTTATGATGATCTTATCAGATTTCTTGAGCCAGCA 960  
QY 880 GCGCTCTTGAATTCAGGAGAGTGGCAGATACAGTATGAGATCTTCTACTGAGGCTTC 939  
DB 961 GCGCTCTTGAATTCAGGAGAGTGGCAGATACAGTATGAGATCTTCTACTGAGGCTTC 1020  
QY 940 TGTGGAAGCTCCAACTCCAACTGATCATCCTGGAGGCGACAAATGGGAGTTCAAGA 999  
DB 1021 TGTGGAAGCTCCAACTCCAACTGATCATCCTGGAGGCGACAAATGGGAGTTCAAGA 1080  
QY 1000 TGACAGACCTGATGAGTGGCTGGGTTGGGAGAGAGAAAGCAACCTTAACATGA 1059  
DB 1081 TGACAGACCTGATGAGTGGCTGGGTTGGGAGAGAGAAAGCAACCTTAACATGA 1140  
QY 1060 ACTATGCAAACTCAGCGCTGCACTTGTGCTACTATGACAAATAATTTATGACTAAG 1119  
DB 1141 ACTATGCAAACTCAGCGCTGCACTTGTGCTACTATGACAAATAATTTATGACTAAG 1200  
QY 1120 TTCTATGTAACGCTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCTCCAGC 1179  
DB 1201 TTCTATGTAACGCTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCTCCAGC 1260  
QY 1180 CTCACCTCCAGATCATCATGATGACAAATACCATAGACCTCCCTACATGATGCTCT 1239  
DB 1261 CTCACCTCCAGATCATCATGATGACAAATACCATAGACCTCCCTACATGATGCTCT 1320  
QY 1240 ACCATGCAACCCCGAGAAATGAACTTTGTAGCTCCCATCCCTGCTTGGCCGCTAA 1299  
DB 1321 ACCATGCAACCCCGAGAAATGAACTTTGTAGCTCCCATCCCTGCTTGGCCGCTAA 1380  
QY 1300 CTTTCATCAGCTTTTGTGCTGCTTAAATCATCTGGAATTCACCAACTGAGAGCATCT 1359  
DB 1381 CTTTCATCAGCTTTTGTGCTGCTTAAATCATCTGGAATTCACCAACTGAGAGCATCT 1440  
QY 1360 ACCCATATACAGGCTGCGAGCTGCTCATATGCTTCCCATCTTGGACCTACTAGTAAG 1419  
DB 1441 ACCCATATACAGGCTGCGAGCTGCTCATATGCTTCCCATCTTGGACCTACTAGTAAG 1500  
QY 1420 TGGGGAAGAAAGAGCGCCAGAGAAA 1447  
DB 1501 TGGGGAAGAAAGAGCGCCAGAGAAA 1528

RESULT 3  
AAO50644  
ID AAO50644 standard; cDNA; 2938 BP.  
XX  
XX AAO50644;  
XX  
XX 26-MAY-1994 (first entry)  
XX  
XX Human Hum-Fil-1 gene clone BM025.  
DB 601 CGCGGAGCTTAAAGCAGATATCTCTCTGACACCTACACTGCTAGAGAGCTCTC 660  
QY 655 ----- 654  
DB 661 TTCACATTTGACTTCAGATGATGTGATAAGGCTTACAAAACCTCTCCAGGTTAATGC 720  
QY 655 -----GAGGAGCCACTTTTATTTTCCAAATATGACTGATTTACCCAGAG 699  
DB 721 ATGCTAGAAACACAGGAGAGCCACTTTTATTTTCCAAATATGACTGATTTACCCAGAG 780  
QY 700 CAACGCAAAAGATTAACAAGAGGCGAGATTACCTTATGAGCAAGCAGAGATCAGCGT 759  
DB 781 CAACGCAAAAGATTAACAAGAGGCGAGATTACCTTATGAGCAAGCAGAGATCAGCGT 840  
QY 760 GAGAGATCACAGGATCCAGCTGATCAAAAGGTACCAACCATCTTCAACAGTGC 819  
DB 841 GAGAGATCACAGGATCCAGCTGATCAAAAGGTACCAACCATCTTCAACAGTGC 900  
QY 820 CCAAAACAGAAAGACACAGCTCTTATGATGATCTTATCAGATTTCTTGAGCCAGCA 879  
DB 901 CCAAAACAGAAAGACACAGCTCTTATGATGATCTTATCAGATTTCTTGAGCCAGCA 960  
QY 880 GCGCTCTTGAATTCAGGAGAGTGGCAGATACAGTATGAGATCTTCTACTGAGGCTTC 939  
DB 961 GCGCTCTTGAATTCAGGAGAGTGGCAGATACAGTATGAGATCTTCTACTGAGGCTTC 1020  
QY 940 TGTGGAAGCTCCAACTCCAACTGATCATCCTGGAGGCGACAAATGGGAGTTCAAGA 999  
DB 1021 TGTGGAAGCTCCAACTCCAACTGATCATCCTGGAGGCGACAAATGGGAGTTCAAGA 1080  
QY 1000 TGACAGACCTGATGAGTGGCTGGGTTGGGAGAGAGAAAGCAACCTTAACATGA 1059  
DB 1081 TGACAGACCTGATGAGTGGCTGGGTTGGGAGAGAGAAAGCAACCTTAACATGA 1140  
QY 1060 ACTATGCAAACTCAGCGCTGCACTTGTGCTACTATGACAAATAATTTATGACTAAG 1119  
DB 1141 ACTATGCAAACTCAGCGCTGCACTTGTGCTACTATGACAAATAATTTATGACTAAG 1200  
QY 1120 TTCTATGTAACGCTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCTCCAGC 1179  
DB 1201 TTCTATGTAACGCTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCTCCAGC 1260  
QY 1180 CTCACCTCCAGATCATCATGATGACAAATACCATAGACCTCCCTACATGATGCTCT 1239  
DB 1261 CTCACCTCCAGATCATCATGATGACAAATACCATAGACCTCCCTACATGATGCTCT 1320  
QY 1240 ACCATGCAACCCCGAGAAATGAACTTTGTAGCTCCCATCCCTGCTTGGCCGCTAA 1299  
DB 1321 ACCATGCAACCCCGAGAAATGAACTTTGTAGCTCCCATCCCTGCTTGGCCGCTAA 1380  
QY 1300 CTTTCATCAGCTTTTGTGCTGCTTAAATCATCTGGAATTCACCAACTGAGAGCATCT 1359  
DB 1381 CTTTCATCAGCTTTTGTGCTGCTTAAATCATCTGGAATTCACCAACTGAGAGCATCT 1440  
QY 1360 ACCCATATACAGGCTGCGAGCTGCTCATATGCTTCCCATCTTGGACCTACTAGTAAG 1419  
DB 1441 ACCCATATACAGGCTGCGAGCTGCTCATATGCTTCCCATCTTGGACCTACTAGTAAG 1500  
QY 1420 TGGGGAAGAAAGAGCGCCAGAGAAA 1447  
DB 1501 TGGGGAAGAAAGAGCGCCAGAGAAA 1528

KW chromosomal translocation; chimeric; chimera; Ewing sarcoma;  
 KW EWS gene; malignant melanoma; hum-flt-1;  
 KW primitive peripheral neuroectodermal tumour; human chromosome 11;  
 KW human chromosome 22; ss.  
 XX  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 143..1501  
 FT /\*tag= a  
 FT /product= HUM-FLT-1  
 FT polyA-signal 2908..2913  
 FT /\*tag= b  
 XX MO9323549-A.  
 XX  
 XX 25-NOV-1993.  
 XX  
 XX 19-MAY-1993; 93WO-FR00494.  
 XX  
 XX 20-MAY-1992; 92FR-0006123.  
 XX  
 XX (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 XX Aurias A, Delattre O, Desmaza C, Melot T, Peter M;  
 XX Plougastel B, Thomas G, Zucman J;  
 XX  
 DR WPI: 1993-386580/48.  
 DR P-PSDB; AAR44556.  
 XX  
 PT New nucleic acid of EWS gene and its hybrid(s) - contg. gene  
 PT sequence involved in chromosomal trans-location, also derived  
 PT of mRNA, probes, fusion proteins etc., for diagnosis and treatment  
 PT of Ewing sarcoma and melanoma  
 XX  
 PS Disclosure; Fig 7; 123bp; French.  
 CC The probe 11R1 was used to screen a human marrow cDNA library  
 CC (Clontech cat. # HL1058). The clone BM025 was identified and  
 CC sequenced. It represents the entire coding region together with  
 CC 5'- and 3'-UTRs of the Hum-Flt-1 gene.  
 XX  
 SQ Sequence 2938 BP; 847 A; 692 C; 675 G; 724 T; 0 other;  
 Query Match 39.0%; Score 563.8; DB 14; Length 2938;  
 Best Local Similarity 66.5%; Pred. No. 4.8e-175;  
 Matches 913; Conservative 0; Mismatches 427; Indels 33; Gaps 6;

494 ACCGACACGAGAGAGATCATCTGCTCCCGCAGACCCACACAGCTGTGACACAGAGCAT 553  
 477 GTACGGCAGTGGCGGAGTGGGACGTGAGAGATGTCTTCCAGCGTGCATCTTG 536  
 554 GTGAGGCAATGGCTGGAGTGGCCATTAAGAGATATGCTTGATGGAGATGACATCTC 613  
 537 TTGTTCCAGAACTTGTATGGGAAAGAGTGTGTAATATGACCAAGATGACTTCCAGAGA 596  
 614 TTTTTCAGAACTATGATGAGGCAAGGACTGTGTAATATGAAACAGAGAGACTCTCTCCG 673  
 597 CTCACGCCGAGCTTAAACGAGATATCTCTCTGACACCTTACACTACCTCAGAGAGA 656  
 674 GCCACCACTCTTAAACAGGAGAGTGTCTTGTACACCTCAGTACTCAGAGAGAG- 732  
 657 GGAGCCACTTTTATTTTCCAAATACATCAGTTTACCAGAGACCAAGCAAGATACAGA 716  
 733 -----TTCACTGCTGGCTTATATACACTCCACACGACCAATCTTCACTGATGAT 787  
 717 ACAAGGCCAGATTTTACCTTATGAGCAAGCAGAGATCAGCTGAGAGTCAAGCCAT 776  
 788 GTCAAGAGAGACCTTCTTATGACTGATCAGAAAGAGACATGGGCAATTAACATGAT 847  
 777 CC---CACTCAGTCAAAAGCTACCCCAACATCATCTTCAACAGTCCCAAAACAGAAAG 833  
 848 TCTGCTCTCAACAAAGTCTCTCCCTTGGAGGGGACAAACGATCACTAAGATATACAG 907  
 834 CAGGCTCTCAGTATGATCTTATACAGATCTTGGACCGACGACGCGCTTGCAAT 893  
 908 CAAGGCGCCAGCAGATCTCTTATGATCAGATCTTGGGCGCCGACGACGATCTCTTAC 967  
 894 CAGAGAGTGGGAGATACAGTATGAGCAGTCTTCTACTGAGGCTGTGTGACAGCTCC 953  
 968 CTGGAAGGGGAGATACAGCTGTGAGCATCTCTCTGAGAGTGTCTCTCCAGAGGCC 1027  
 954 AACTCCACTGCTACCTTGGGAGGACAAATGGGAGTTCAAGATGACAGACCTGAT 1013  
 1028 AAGCGCAGCTGTATACCTTGGGAGGAGGACCAAGGAGTCAAAATGACGAGCCCGAT 1087  
 1014 GAAGTGGCTGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1073  
 1088 GAGTGGCCAGGCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1147  
 1074 AGCCGTGACTTGTGCTATGATGACCAAAATATATGATGATGATGATGATGATGATG 1133  
 1148 AGCGGGGCGCTCGTATATGATGATGATGATGATGATGATGATGATGATGATGATG 1207  
 1134 TATGCTACAAATTTGATTTCCAGGAGATGCTCAGAGCCCTCCAGGCTCAGCTCAGAA 1193  
 1208 TATGCTTACAAATTTGATTTCCAGGAGATGCTCAGAGCCCTCCAGGCTCAGGACGAG 1267  
 1194 TCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1253  
 1268 TCGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1327  
 1254 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1313  
 1328 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1387  
 1314 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1371  
 1388 TTTGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1447  
 1372 -----GGCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1417  
 1448 GTCCCGCGCATCTTCAACACCAAGTGTCTTCACTTACAGTACTACTACTA 1500

RESULT 4  
 ABR84139  
 ID ABR84139 standard; cDNA; 2957 BP.  
 XX  
 AC ABR84139;

14-AUG-2002 (first entry)  
Human CDNA differentially expressed in granulocytic cells #710.  
Human; ss: granulocytic cell; DNA chip; bacterial infection;  
viral infection; parasitic infection; protozoal infection;  
fungal infection; sterile inflammatory disease; psoriasis;  
rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
cardiac reperfusion injury; renal reperfusion injury; ARDS;  
adult respiratory distress syndrome; inflammatory bowel disease;  
Crohn's disease; ulcerative colitis; periodontal disease;  
granulocyte activation; chronic inflammation; allergy.  
Homo sapiens.  
W0200228999-A2.  
11-APR-2002.  
03-OCT-2001; 2001WO-US30821.  
03-OCT-2000; 2000US-237189P.  
(GENE-) GENE LOGIC INC.  
Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
WPI; 2002-435328/46.  
Detecting granulocyte activation by detecting differential expression  
of genes associated with granulocyte activation, which serves as  
diagnostic markers that is useful for monitoring disease states and  
drug toxicity  
Claim 1; SEQ ID NO 710; 114pp; English.  
The invention relates to detecting (M1) granulocyte (GC) activation  
(GCA), by detecting the level of expression of gene(s) (Gs) identified by  
DNA chip analysis as given in the specification, and comparing  
the expression level to an expression level in an unactivated  
GC, where differential expression of Gs is indicative of GCA.  
Also included are modulating (M2) GA by contacting GC with an agent  
that alters the expression of at least one gene in Gs; (2) screening (M3)  
for an agent capable of modulating GCA or an inflammation (especially  
chronic) in a tissue, an allergic response in a subject, exposure of a  
subject to a pathogen or sterile inflammatory disease using the  
gene expression profile; (3) detecting (M4) an inflammation (especially  
chronic) in a tissue, an allergic response in a subject, exposure of a  
subject to a pathogen or sterile inflammatory disease, by detecting the  
level of expression in a sample of the tissue of gene(s) from Gs, where  
the level of expression of the gene is indicative of inflammation;  
(4) treating (M5) an inflammation (especially chronic) or in a tissue,  
an allergic response in a subject, exposure of a subject to a pathogen  
or sterile inflammatory disease, by contacting a tissue having  
inflammation with an agent that modulates the expression of gene(s)  
from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for  
modulating GA; M3 is useful for screening an agent capable of modulating  
GCA preferably in an inflammation in a tissue; M4 is useful for  
detecting an inflammation (especially chronic) in a tissue, an allergic  
response in a subject, exposure of a subject to a pathogen or sterile  
inflammatory disease (e.g. psoriasis, rheumatoid arthritis,  
glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal  
reperfusion injury, ARDS, adult respiratory distress syndrome,  
inflammatory bowel disease, Crohn's disease, ulcerative colitis,  
periodontal disease; also bacterial infection, viral infection,  
parasitic infection, protozoal infection, fungal infection and M5 is  
useful for treating one of the above conditions. The present  
sequence represents a gene differentially expressed in granulocytes.  
Note: The sequence data for this patent did not form part  
of the printed specification, but was obtained in electronic  
format directly from WIPO at  
ftp.wipo.int/pub/published\_pcl\_sequences.

XX	Sequence	2957 BP, 855 A, 698 C, 680 G, 724 T, 0 other:
sq	Query Match	38.9%; Score 562.2; DB 24; Length 2957;
	Best Local Similarity	66.4%; Pred. No. 1.6e-174;
	Matches	912; Conservative 0; Mismatches 428; Indels 33; Gaps 6;
QY	ATGCGACGACATTATTAAGGAAGCATTTATCAGTGTGAGTGAAGACGACTCTTGTGTGAG	122
DB		173
QY	ATGACGGGAGCATTTAAAGAGGCTCTGTGGGTGTGACGACGACGACTCTCTTTGTAC	232
DB		123
QY	TGTGGCTACGGATG-----CCGACCTGTCCAAGACAGAAATGAGACGCTCCCTCCAGT	179
DB		233
QY	TCAGGTACGGAGCGGCGACGCCATCTCCCAAGGCGCATGACTGCTCCGGGAGTCC	292
DB		180
QY	GAATATGAGGCAACATCAAAAGATGAGCCCGGCGCTTCCCGACGAGAGCTGTTATACAG	239
DB		293
QY	GACTATGSGGACGCCCCACACATATCAACCCCTCCACACAGCAGAGATGGATCATATCAG	352
DB		240
QY	CCCCCGGCGCAGATTACCATTTAAGTGAAGTGTAAACCAACGAGTTTAATGGCTCAAG	299
DB		353
QY	CC-----AGTGAAGGTCAACGTCAAGCGGGGAGTGTGACACATGAATGATCCAG	403
DB		300
QY	AATTCACCTGATGATGTGACGCGTGGCCAAAGAGGGAATATGTATGACATTCACATAT	359
DB		404
QY	GAGTCTCGGAGACTGTGACGGTTTAAATGACACAGCTGTGTGGGCGGAGCCGATCC	463
DB		360
QY	GTTGGATGACTATGGAAGCTACATCTGGAAGAAGACAT-----ATTCCGCTCCAAATATG	416
DB		464
QY	AACCCATGAATACACACAGCTATATGACGAGAAATGGCCCCCTCTCTCCCAACATG	523
DB		417
QY	ACAACCAATGAAACGAAGATTATGTGTGCACAGATCCCTACGTTATGAGACAGACCAT	476
DB		524
QY	ACCACACAGAGAGAGAGATCATGTGCTCCCGCAGACCCACACATGTGTGACACAGAGCAT	583
DB		477
QY	GTAGGCGAGTGGCTGTGAGTGGGCGAGTGAAGGATGTGTTCCAGACGTGGACATCTTG	536
DB		584
QY	GTAAGGCAATGGCTGTGAGTGGGCCATAAAGAGTACACATTTGATGGAATGACACATCC	643
DB		537
QY	TTGTTCCGAAACATTGAGGGGAAGAGTGTGTAAATGACCAAAAGATGATCTCCAGAGA	596
DB		644
QY	TTTTTCCGAACATGTGATGGCAGAGAACTGTGTAAATGAACAGAGAGACTTCTCCGC	703
DB		597
QY	CTCAAGCCGAGCTATTAACGAGATATCTCTCTGTACACACTACATCTCAGAGAGAGA	656
DB		704
QY	GCCACACACCTCTACAAACAGGAAAGTGTGTGTGCACACCTCAGTTACTCAGGGAAG-	762
DB		657
QY	GGAGCCACTTTATTTTCCAAATATCATCAGTTTAAACCAAGACCAAGCAAGATTAACA	716
DB		763
QY	-----TTCACGTGCGCTTATATACAACTCCACACGACCAATCTCCAGATTGAGT	817
DB		717
QY	ACAAGGCCAGATTTACCTTATGAGACACGAGAGAGATCAGGTGGAGAGTCAACGCAAT	776
DB		818
QY	GTCAAGAAGACCTCTTATATGACTCATGTCAAGAAAGAGAGCTTGGGCAATTAATCATGAT	877
DB		777
QY	CC-----CACTCAGTCAAAAAGCTACCCACATCATCTTTAACAGTGGCCCAAAACAGAAAGC	833
DB		878
QY	TCTGGCCCTCAACAAAAGTCTCTCCCTTGTGAGGGGACAAAGATCAGTAAATAATACAGAG	937
DB		834
QY	CAGGCTCTCAGTTAGATCTTTATCAGATTTCTTGTGACCGGACGACGACGCGTCTTGCANA	893
DB		938
QY	CAAGGGCCCCAGCAGATCCGTATACAGATCTGTGGGCCGACACAGAGTGGCTTATGAC	997
DB		894
QY	CCAGGAGATGGGCGAGATACAGCTATATGACAGTTCTCTACTGAGAGCTTCTGTGGACAGTCC	953
DB		998
QY	CTTGAAGCGGGCGAGATTCACGCTGTGGCAATTTCTCTGTGAGCTGTCTTCGACAGGCC	1057
DB		954
QY	AATTCACAACTCATACACCTGTGGAGGGCCAAATGGGGATCCAGATATACAGACCTCAT	1013
DB		1058
QY	AAAGCCAGCTATATCAGCTGTGGAGGGGACCAAGCGGAGATTCAAATATACGAGACCCGAT	1117
DB		1014
QY	GAGTGGCTGGGCTTGGGAGAGAGAAAAGCAAACTTAACATGAACTATGACAAACTC	1073
DB		

Db	1118	GAGTGGCCAGGGCGCTGGGGGAGCGGAAAAAGCAAGCCCAACATGAAATTAGACAACTG	1177
QY	1074	AGCGGTGCACTTCGCTACACTACTATGACAAAAAATATTATGACTAAAGTTTCATGTAAAGC	1133
Db	1178	AGCGGGGCGCCCTCCCTTATTTACTATGATAAAAAACATTATTGACCAAAAGTGACAGCGCAANA	1237
QY	1134	TATGCGTACAAATTTGATTTCCACAGGAATGCGTAGGCGCCCGACGCTCAGCCACCCTCAGAA	1193
Db	1238	TATGCTTACAAATTTGACTTCCACAGGAGGATGGCCAGGCTCTGCAGCCACATCCGACGAG	1297
QY	1194	TCATCCATGTACAAATTAACCATCCAGACCTCCCTACATAGATTTCTTACATGACACGCC	1253
Db	1298	TCGTCATGTACAAAGTAGTACCCCTTTCGACATCTCTACATGCGCTTCTTACATGCGCCACGAG	1357
QY	1254	CAGAAGATGAACCTTGTAGCTCCCCATCCCGCTGTTGGCCCGTAACCTCATCCAGGCTT	1313
Db	1358	CAGAAGTGAACCTTGTCCCTCCCATTCATCTCCATGCGCTGTCACTTCTCCAGGCTTC	1417
QY	1314	TTTTCGTGCGCCCTAATTCATACACTGGAATTCACCAACTGGAGGCACTCTACCCCAATACA--	1371
Db	1418	TTTTCGAGCGCGCATCACAACTACTGGACCTCCCGCCAGGGGGGAATCTACCCCAACCCCAAC	1477
QY	1372	-----GGGTGCGCAGCGTGCATATACCTTTCCCATCTGTGGACACTCTACTACTA	1417
Db	1478	GTCCCCCGCATCTTAACACCAAGTGCCTTTCACACTTAGGACACTCTACTACTA	1530

RESULT 5  
AAH02915  
ID AAH02915 standard: DNA: 2954 BP

AC AAH02915;

DT 15-JUN-2001 (first entry)

Human shear stress-response coding sequence SEQ ID NO: 83.

human; shear stress-response protein; vascular disease

OS Homo sapiens.

PN WO200125427-A1

PD 12-APR-2001.

PF 02-OCT-2000; 2000WO-JP06840.  
yy

PK 01-001-1999; 99JP-0280976;  
XX XX

PA (NOJI/) NOJIMA H.

PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

DR P-PSDB; AAB90792.  
YY

PI DNA sequences, proteins encoded by them and antibodies against them useful in diagnosis and treatment of vascular disease caused by PT

XX  
XX  
XX

XX

CC number of human shear stress response proteins. These are useful in the

CC arteriosclerosis, including heart failure, post-PTCA restenosis and

XX  
XX

Query Match	37.6%;	Score 544.6;	DB 22;	Length 2954;
Best Local Similarity	-66.1%;	Pred. No. 1.1e-168;		
Matches 908; Conservative	0;	Mismatches 429;	Indels 36;	Gaps 7;

QY	63	ATGGCAACACCTATTTAAGGAAGCATTTATCAGTGGAGTGTGAAGACCACTCTCTGTTTATG	122
Db	173	ATGACGGGACATATTAAGAGGCTCTGTGGGTGGACGACGACCACTCCCTTTTGAC	232
QY	123	TGTGCTACGGATCG --- CCCCACTTCCAAAGACAGAAATGACAGCCTCTCTTCCAGT	179
Db	233	TCAGGTATGAGGCGCGCAGCCCTATCTCCCAAGGCGCATGATCACTGCTCCGGAGTCTT	292
QY	180	GAATATGGCAACATCTAAAGATGAGCCCGCGCTTCCCAAGCAGAGATGGTTATACAG	239
Db	293	GACTACGGGACGCCCCACAGATCAACCCCTCCACACGACGAGATGGATCAATACG	352
QY	240	CCCCCGGCGAGATTACCATTAAGATGAGTGTAAACCAACCAGGTTAATGGGTCAAG	299
Db	353	CC-----AGTGGAGGTCAACGTCGAAGGGGGAGTATGACACATCAATGAATGATCCAG	403
QY	300	AATTCACCTGATGATCGAGGCTGGCCAAAGAGGAAATGGTATGAGAGTTACAGACAT	359
Db	404	GAGTCTCGGTGAGTGTGACGGTTTACCAATATGACCAACGTGTGGTGGGAGCGCATCC	463
QY	360	GTTGGGATGAATATGAGAGCTACATGGAAGAGAAGCAT---ATTCCGCTCCAAATATG	416
Db	464	AAACCCATGAATCTACACAGCATATATGACGAGAAATGGCCCCCTCTCCCAACATG	523
QY	417	ACAACTGAAACGAAGTATTGTGCCAGACATCTTACGTTATGGACGACAGCAT	476
Db	524	ACCACCAACGAGAGAGATGATCTGCCGACAGCCCAACATGTGTGACACAGAGAT	583
QY	477	GTACGCGAGTGTGCTGAGTGGGCGAGTGAAGGAGTATGTCTTCCGACGTGCAATCTTG	536
Db	584	GTGAGGCATATGGCTGTGAGTGGGCCATTAAGGAGTACAGCTTGATGGAGATGACACATCC	643
QY	537	TTGTTCAGAACTTATGATGGGAAGAGTGTGTAAATGACCAAAAGATGACTTCCAGAGA	596
Db	644	TTTTTTCAGAACTGATGATGGCAAGAACTGTAAATGAAACAAAGAGATCTTCTCCGC	703
QY	597	CTCAAGCCGACCTATTAACGCAGATATCTCTGTGCACACCTACACTACTCTAGAGAGAGA	656
Db	704	GCCACACACCTCTCAACACGGAAGTGTGTGTGCACACCTCACTTACCTTAGGGAAG-	762
QY	657	GGAGCCACTTTATTTTCCAAATACATCATATTACCAGAGAAGCAACGAATAAACA	716
Db	763	-----TTCACTGTGTGGCTATTAATATCAACCTCCACACCCGACCAATCTCTACAGTTAGT	817
QY	717	ACAAGGCCAGATTACCTTATGACCAAGCAGAGAGATCAGCGTGGACGATACAGCCAT	776
Db	818	GTCAAGAGACACCTTCTTTATGACTCATGTCAGAAAGAGAGACCTTGGGGCAATATCATGAT	877
QY	777	CC-----CACTAGTGTAAAGCTAACCCAAACCATCTTCAACAGTGGCCCAAAACGAAGAC	833
Db	878	TTGTGGCTCAACAAAAGTCTCCCTTGGAGGGGACAAACGATCATGTAACATATACAG	937
QY	834	CAGCGTCTCAGTTAGATCTTATTCAGATTTTGGACGACAGCAGCCGCTTGCAAAT	893
Db	938	CAAGGGCCCCAGCAGATCCGTATCAAGTCTCTGGGGCCGACACAGATGCGCTTAGCAAC	997
QY	894	CCAGGAGATGGGCAGATACAGCTTATGGCAGTTCTACTGAGACTTGTGTGGACAGCTCC	953
Db	998	CTGTGAAGCGGGCAGATCTACAGCTGTGGCAATCTCTCGGAGCTGTCTCCGACAGGCC	1057
QY	954	AACATCAACTGATCACCTGGGAGGGGCAAAATGGGGATCAAGATGACAGACCCGAT	1013
Db	1058	AACGCCAGCTATACCTGGGAGGGGACCAACGGGGGATTTCAAAATGACGACCCCGAT	1117
QY	1014	GAACTGTGCGCGCTGGGAGAGAGAAAAGCAAACTTACATGAACTATGACAAACTC	1073
Db	1118	GAGGTGGCAGGCGCTGGGGGAGCGGAGAAAAGCAAGCCCAACATGATTAATACGACAGCTG	1177
QY	1074	AGCCGTGCACTTGCCTACTACTATGACAAAAATATATGACTTAAAGTTGATGTAACGC	1133



XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483447/52.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human fetal liver -  
 XX  
 PS Claim 4; SEQ ID NO 14308; 639bp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at fip.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;  
 XX  
 Query Match 26.7%; Score 387; DB 22; Length 567;  
 Best Local Similarity 83.7%; Pred. No. 5.2e-117;  
 Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
 XX  
 QY 896 AGGAGTGGGAGATACAGTATGCAAGTCTCTACTGAGCTTCTGTGGACAGCTCCAA 955  
 DB 523 AGGCAGTGGGAGATACAGTATGCAAGTCTCTCTGAGCTCTGTGGAGACAGCTCAA 464  
 QY 956 CTCACATGCTACCTGAGGAGGAGCAAAATGGGAGTTCAAGATGACACCCCTGATGA 1015  
 DB 463 CTCACATGCTACCTGAGGAGGAGCAAAATGGGAGTTCAAGATGACACCCCTGATGA 404  
 QY 1016 AGTGGCTGGGCTTGGGAGAGAGAAAGCAAACTTAACATGATGATGACAACTCAG 1075  
 DB 403 GGTGGCCGGCGCTGGGAGAGAGCAAAAGCAAACTTAACATGATGATGACAACTCAG 344  
 QY 1076 CGGTGACCTTGGCTACTACTATGACAAATAATTAATGACTTAAGATTGATGATGATG 1135  
 DB 343 CGGTGACCTTGGCTACTACTATGACAAATAATTAATGACTTAAGATTGATGATGATG 284  
 QY 1136 TGGTCAAAATTTGATTTCCAGGAAATCGTCAAGCCCTCCAGCTCACCCTCCAGAAATC 1195  
 DB 283 CGGTGACCTTGGCTACTACTATGACAAATAATTAATGACTTAAGATTGATGATGATG 224  
 QY 1196 ATTCAGTATGCAATATACCATACAGACTCCCTACTAGATGATGATGATGATGATGATG 1255  
 DB 223 ATTCAGTATGCAATATACCATACAGACTCCCTACTAGATGATGATGATGATGATGATG 164  
 QY 1256 GAAGATGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1315  
 DB 163 GAAGATGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 104  
 QY 1316 TGGTGGCCCTTAATCCATACCTGGAATTCACCAATGAGGATGATGATGATGATGATGATG 1375  
 DB 103 TGGTGGCCCTTAATCCATACCTGGAATTCACCAATGAGGATGATGATGATGATGATGATG 44  
 QY 1376 GGCAGCTGCTATATGCTTCCATCTTGAGCACTTACTACTAA 1418  
 DB 43 CCCGACAGCAGCATATGCTTCTCATCTGAGCACTTACTACTAA 1

RESULT 8  
 ID ABA33090 standard; DNA; 567 BP.  
 XX  
 AC ABA33090;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Probe #11556 for gene expression analysis in human heart cell sample.  
 XX  
 KW Human; gene expression; heart; microarray; vascular system; probe;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001MO-US00655.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-48899/53.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts -  
 XX  
 PS Claim 4; SEQ ID NO 11556; 530bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart. The  
 CC present sequence is one such probe. The probes may be used for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from the human heart via microarrays. By measuring gene expression, the  
 CC probes are useful for predicting, diagnosing, grading, staging,  
 CC monitoring and prognosing diseases of the human heart and vascular system  
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
 CC congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at fip.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;  
 XX  
 Query Match 26.7%; Score 387; DB 22; Length 567;  
 Best Local Similarity 83.7%; Pred. No. 5.2e-117;  
 Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
 XX  
 QY 896 AGGAGTGGGAGATACAGTATGCAAGTCTCTACTGAGCTTCTGTGGACAGCTCCAA 955  
 DB 523 AGGCAGTGGGAGATACAGTATGCAAGTCTCTCTGAGCTCTGTGGAGACAGCTCAA 464  
 QY 956 CTCACATGCTACCTGAGGAGGAGCAAAATGGGAGTTCAAGATGACACCCCTGATGA 1015  
 DB 463 CTCACATGCTACCTGAGGAGGAGCAAAATGGGAGTTCAAGATGACACCCCTGATGA 404  
 QY 1016 AGTGGCTGGGCTTGGGAGAGAGAAAGCAAACTTAACATGATGATGATGATGATGATG 1075  
 DB 403 GGTGGCCGGCGCTGGGAGAGAGCAAAAGCAAACTTAACATGATGATGATGATGATGATG 344  
 QY 1076 CGGTGACCTTGGCTACTACTATGACAAATAATTAATGACTTAAGATTGATGATGATGATG 1135



XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488900/53.  
 DR

XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 XX

PS Example 4; SEQ ID NO: 14717; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of  
 CC the probes of the invention.

XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Query Match 26.7%; Score 387; DB 22; Length 567;

Best Local Similarity 83.7%; Pred. No. 5.2e-117;  
 Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 896 AGGAGTGGGACATACAGTATGCGAGTCTCTACTGAGCTTCTGTGGACAGCTCCAA 955  
 Db 523 AGGACAGTGGGACATACAGTATGCGAGTCTCTACTGAGCTCTGTGGACAGCTCCAA 464

QY 956 CTCACATGATCATCCTGGGAGGACCAATGAGGAGTTCAAGATGACACACCTGATGA 1015  
 Db 463 CTCACATGATCATCCTGGGAGGACCAATGAGGAGTTCAAGATGACACACCTGATGA 1015

QY 1016 AGTGGCTCGGCTTGGGAGAGAGAAAGCAACCTACATGATGATGACAACTGAG 1075  
 Db 403 GGTGGGCGGCGCTGGGAGAGAGAGAAAGCAACCTACATGATGATGACAACTGAG 1075

QY 1076 CCGTGCATCTGCTACTACTATGACAAATATTTATGATGATGATGATGATGATGAT 1135  
 Db 343 CCGGCGCTCTCTACTACTATGACAAATATTTATGATGATGATGATGATGATGAT 1135

QY 1136 TGGCTCAAAATTTGATTTCCAGGAATTCGCTCAGGCGCTCCAGCTCCAGCTCCAGATC 1195  
 Db 283 CGCTCAAAATTTGATTTCCAGGAATTCGCTCAGGCGCTCCAGCTCCAGCTCCAGATC 1195

QY 1196 ATCCATGATCAAAATTTGATTTCCAGGAATTCGCTCAGGCGCTCCAGCTCCAGATC 1255  
 Db 223 ATCCATGATCAAAATTTGATTTCCAGGAATTCGCTCAGGCGCTCCAGCTCCAGATC 1255

QY 1256 GAAGATGATCTTGTAGTCCCATGCGCTGCTTGGCCGTAACTGATCCAGCTTTT 1315  
 Db 163 GAAGATGATCTTGTAGTCCCATGCGCTGCTTGGCCGTAACTGATCCAGCTTTT 1315

QY 1316 TGGTGGCCCTTAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375  
 Db 103 TGGTGGCCCTTAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375

QY 1376 GCCAGCTGCTCATATGCTTCCCATCTGGGACCTGACCTACTATA 1418  
 Db 43 GCCAGCTGCTCATATGCTTCCCATCTGGGACCTGACCTACTATA 1418

RESULT 11  
 AAI20932/c  
 ID AAI20932 standard; DNA; 567 BP.  
 XX

XX AAI20932;

XX 12-OCT-2001 (first entry)

DE Probe #10865 for gene expression analysis in human cervical cell sample.

XX Probe: human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer; ss.  
 XX

OS Homo sapiens.

PN W0200157278-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US006070.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

DR Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX

PS Claim 25; SEQ ID NO 10865; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes  
 CC (SENPs). The present sequence is one such probe. The SENPs are derived  
 CC from human Hela cells. The SENPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging  
 CC of diseases of the cervix, notably cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at [http://wipo.int/pub/published\\_pcl\\_sequences](http://wipo.int/pub/published_pcl_sequences).

XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Query Match 26.7%; Score 387; DB 22; Length 567;

Best Local Similarity 83.7%; Pred. No. 5.2e-117;  
 Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 896 AGGAGTGGGACATACAGTATGCGAGTCTCTACTGAGCTTCTGTGGACAGCTCCAA 955  
 Db 523 AGGACAGTGGGACATACAGTATGCGAGTCTCTACTGAGCTTCTGTGGACAGCTCCAA 464

QY 956 CTCACATGATCATCCTGGGAGGACCAATGAGGAGTTCAAGATGACACACCTGATGA 1015  
 Db 463 CTCACATGATCATCCTGGGAGGACCAATGAGGAGTTCAAGATGACACACCTGATGA 1015

QY 1016 AGTGGCTCGGCTTGGGAGAGAGAAAGCAACCTACATGATGATGATGATGATGATGAT 1075  
 Db 403 GGTGGGCGGCGCTGGGAGAGAGAGAAAGCAACCTACATGATGATGATGATGATGATGAT 1075

QY 1076 CCGTGCATCTGCTACTACTATGACAAATATTTATGATGATGATGATGATGATGATGAT 1135  
 Db 343 CCGGCGCTCTCTACTACTATGACAAATATTTATGATGATGATGATGATGATGATGAT 1135

QY 1136 TGGCTCAAAATTTGATTTCCAGGAATTCGCTCAGGCGCTCCAGCTCCAGCTCCAGATC 1195  
 Db 283 CGCTCAAAATTTGATTTCCAGGAATTCGCTCAGGCGCTCCAGCTCCAGCTCCAGATC 1195

QY 1196 ATCCATGATCAAAATTTGATTTCCAGGAATTCGCTCAGGCGCTCCAGCTCCAGATC 1255  
 Db 223 ATCCATGATCAAAATTTGATTTCCAGGAATTCGCTCAGGCGCTCCAGCTCCAGATC 1255

QY 1256 GAAGATGATCTTGTAGTCCCATGCGCTGCTTGGCCGTAACTGATCCAGCTTTT 1315  
 Db 163 GAAGATGATCTTGTAGTCCCATGCGCTGCTTGGCCGTAACTGATCCAGCTTTT 1315

QY 1316 TGGTGGCCCTTAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375  
 Db 103 TGGTGGCCCTTAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375



||||| 1376 GCCAGCTGCTCATATGCTTCCCATCTTGGACACTACTACTA 1418  
||| 43 CCCACACGACCATATGCTTCTCATCTGGGACACTTACTACTA 1  
RESULT 12  
AAI46176/c  
ID AAI46176 standard; DNA; 567 BP.  
XX  
AC AAI46176;  
XX  
DE 17-OCT-2001 (first entry)  
XX  
DE Probe #14862 used to measure gene expression in human placenta sample.  
XX  
KM Probe; microarray; human; placenta; antenatal diagnosis;  
KM genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488897/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human placenta -  
XX  
PS Claim 25; SEQ ID No 14862; 654pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENPs).  
CC The present sequence is one such probe. The probes are useful for  
CC producing a microarray for predicting, measuring and displaying gene  
CC expression in samples derived from human placenta. The probes are useful  
CC for antenatal diagnosis of human genetic disorders.  
XX  
SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;  
Query Match 26.7%; Score 387; DB 22; Length 567;  
Best Local Similarity 83.7%; Pred. No. 5.2e-117;  
Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
OY 896 AGGAGTGGGACAGATACATGCTATGTCAGTTCCTCTGAGGCTTGTCGGACACTCCAA 955  
DB 523 AGCGATGGCCAGATCCAGCTTTGGCACTTCTCTGAGGCTCTGTCGGACACTCCAA 464  
OY 956 CTCCAACTGCATCACTTGGAGGACCAAAATGGGAGTTCAAGATGACAGACCTGATGA 1015  
DB 463 CTCAGCTGCATCACTTGGAGGACCAAGGAGGAGTTCAAGATGACAGATCCGACGA 404  
OY 1016 AGTGGCTGGGCTGGGAGAGAGAAAGCAAACTTACATGACATGACAAACTGAG 1075  
DB 403 GGTGGCCCGGCTGGGAGAGCGGAAGCAAACTTACATGACATGACATGAGCTCAG 344  
OY 1076 CCGTCACTTGGCTACTACTATGACAAAATATATTATGACTAACTGATGTAACGCTA 1135

||||| 1376 GCCAGCTGCTCATATGCTTCCCATCTTGGACACTACTACTA 1418  
||| 43 CCCACACGACCATATGCTTCTCATCTGGGACACTTACTACTA 1  
RESULT 13  
AAI06643/c  
ID AAI06643 standard; DNA; 567 BP.  
XX  
AC AAI06643;  
XX  
DE 09-OCT-2001 (first entry)  
XX  
DE Probe #6634 used to measure gene expression in human breast sample.  
XX  
DE Probe; human; breast disease; breast cancer; development disorder; ss;  
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US000661.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-476286/51.  
XX  
PT Novel single exon nucleic acid probe used to measuring gene expression  
XX in a human breast -  
XX  
PS Claim 25; SEQ ID No 6634; 322pp; English.  
XX  
CC The present invention relates to novel single exon nucleic acid probes.  
CC The present sequence is one such probe. The probes are useful for -  
CC measuring human gene expression in a human breast sample, where the probe  
CC hybridises at high stringency to a nucleic acid expressed in the human  
CC breast. The probes are useful for predicting, diagnosing, grading,  
CC staging, monitoring and prognosing diseases of the human breast,  
CC particularly those diseases with polygenic aetiology. The diseases  
CC include: breast cancer, disorders of development, inflammatory diseases  
CC of the breast, fibrocystic changes, proliferative breast disease and

CC non-carcinoma tumours.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other:

Query Match 26.7%; Score 387; DB 22; Length 567;  
 Best Local Similarity 83.7%; Pred. No. 5.2e-117;  
 Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY AGGAGTGGGAGATACAGTATGAGCTTCTACTGAGCTTCTGTGGACAGCTCCAA 955  
 DB 523 AGGAGTGGGAGATACAGTATGAGCTTCTACTGAGCTTCTGTGGACAGCTCCAA 464  
 QY 956 CTCACATGATACCTGGGAGGAGCAATGAGGAGTTCACATGACACACCTGATGA 1015  
 DB 463 CTCACATGATACCTGGGAGGAGCAATGAGGAGTTCACATGACACACCTGATGA 404  
 QY 1016 AGTGGCTCGGCTGTGGGAGAGAGAAACCAACCTAATGATGACAAACCTCAG 1075  
 DB 403 GGTGGCCCGGCGCTGGGAGAGAGAGAAACCAACCTAATGATGACAAACCTCAG 344  
 QY 1076 CCGTGCACCTGCTACTACTATGACAAAATATTATGACTAAAGTTGATGAAAGCTTA 1135  
 DB 343 CCGGCGCCCTCGTACTACTATGACAAAATATTATGACTAAAGTTGATGAAAGCTTA 284  
 QY 1136 TGGCTCAAAATTTGATTTCCAGGAAATCGCTCAGAGCCCTCAGAGCTCCAGAAATC 1195  
 DB 283 CCGCTCAAAATTTGATTTCCAGGAAATCGCTCAGAGCCCTCAGAGCTCCAGAAATC 224  
 QY 1196 ATCCATGACAAATATCCATCAGACCTCCCTTACATGATGATTCCTACATGACACCCCA 1255  
 DB 223 ATCTGTGTAACATATACCTCCCTCAGACCTCCCTTACATGATGATTCCTACATGACACCCCA 164  
 QY 1256 GAGATGAATTTGTGATGCTCCCATCCCTGCTTGGCCGTGATCCATCCAGCTTTT 1315  
 DB 163 GAGATGAATTTGTGATGCTCCCATCCCTGCTTGGCCGTGATCCATCCAGCTTTT 104  
 QY 1316 TGGTGGCCCTTAATCAGTATGAAATTCACCACTGAGAGGATGATCAGATCAGAGCT 1375  
 DB 103 TGGTGGCCCTTAATCAGTATGAAATTCACCACTGAGAGGATGATCAGATCAGAGCT 44  
 QY 1376 GCCAGCTGCTATATGCTTCCATCTTGGACCTTACTAA 1418  
 DB 43 CCCACACGACATATGCTTCTCATCTGCGCACCTTACTAA 1

# RESULT 14

ABSI4215/C  
 ID ABSI4215 standard; DNA; 567 BP.

XX ABSI4215;

DT 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe ORF from lung SEQ ID No 14206.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
 XX chronic obstructive pulmonary disease; interstitial lung disease;  
 XX familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;  
 XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 XX primary ciliary dyskinesia; pulmonary hypertension;  
 XX hyaline membrane disease; open reading frame; ORF.

XX Homo sapiens.

PN WO200186003-A2.

PD 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.  
 PF 04-FEB-2000; 2000US-180312P.  
 XX 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PT Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to  
 measure gene expression in human lung samples -

PS Claim 4; SEQ ID No 14206; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridise at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression to a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a single exon  
 CC probe open reading frame of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

QY Query Match 26.7%; Score 387; DB 24; Length 567;  
 Best Local Similarity 83.7%; Pred. No. 5.2e-117;

Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 896 AGGAGTGGGAGATACAGTATGAGCTTCTACTGAGCTTCTGTGGACAGCTCCAA 955  
 DB 523 AGGAGTGGGAGATACAGTATGAGCTTCTACTGAGCTTCTGTGGACAGCTCCAA 464

XX	ABAS0193; standard; DNA; 473 BP.
XX	ABAS0193;
XX	01-FEB-2002 (first entry)
XX	Human breast cell single exon nucleic acid probe #8888.
XX	Human; microarray; single exon probe; gene expression; breast;
XX	disease; cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200157271-A2.
PD	
XX	09-AUG-2001.
XX	
XX	30-JAN-2001; 2001WO-US00062.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
PA	
XX	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
DR	WPL; 2001-496933/54.
XX	
XX	New spatially-addressable set of single exon nucleic acid probes;
XX	useful for measuring gene expression in sample derived from human
XX	breast, comprises number of single exon nucleic acid probes -
PS	Claim 4; SEQ ID NO 8888; 327pp + sequence listing; English.

[illegible]

Claim 4: SEQ ID NO 8888; 327bp + sequence listing; English.  
New spatially-addressable set of single exon nucleic acid probes;  
useful for measuring gene expression in sample derived from human  
breast, comprises number of single exon nucleic acid probes -

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 10:00:14 : Search time 91.3059 Seconds

(Without alignments)

4860.160 Million cell updates/sec

Title: US-09-902-772-1

Perfect score: 1447

Sequence: 1 gaattccgcgaacgaataat.....gaaagaagcgccaagaataa 1447

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCrus.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1447	100.0	1447	US-08-878-177-1	Sequence 1, Appl
2	1356	93.7	1528	US-08-878-177-3	Sequence 3, Appl
3	563.8	39.0	2938	US-08-343-443B-3	Sequence 3, Appl
4	180	12.4	1752	US-09-360-779-1	Sequence 1, Appl
5	180	12.4	1752	US-09-435-335-1	Sequence 1, Appl
6	124.6	8.6	1604	US-08-306-691B-43	Sequence 1, Appl
7	124.6	8.6	1604	PCT-US93-06251-9	Sequence 4, Appl
8	122.2	8.4	2268	US-09-344-579-1	Sequence 9, Appl
9	117.4	8.1	2667	US-08-469-412A-1	Sequence 1, Appl
10	117.4	8.1	2667	US-09-021-715-1	Sequence 1, Appl
11	115.2	8.0	328	US-08-343-443B-5	Sequence 5, Appl
12	96	6.6	2064	US-08-875-944B-1	Sequence 1, Appl
13	96	6.6	2064	US-09-116-049-3	Sequence 3, Appl
14	95.2	6.6	2410	US-08-780-835B-1	Sequence 1, Appl
15	95.2	6.6	2410	US-09-303-268-1	Sequence 1, Appl
16	95.2	6.6	2410	US-09-116-049-1	Sequence 1, Appl
17	88.6	6.1	2266	US-09-213-767-1	Sequence 1, Appl
18	87.4	6.0	2544	US-08-469-412A-6	Sequence 1, Appl
19	87.4	6.0	2544	US-09-021-715-6	Sequence 6, Appl
20	67.6	4.7	65042	US-09-784-316-3	Sequence 3, Appl
21	59.8	4.1	852	US-09-020-956-44	Sequence 44, Appl
22	59.8	4.1	852	US-09-030-607-44	Sequence 44, Appl
23	59.8	4.1	852	US-09-605-785-44	Sequence 44, Appl
24	59.8	4.1	852	US-09-439-313-44	Sequence 44, Appl
25	59.8	4.1	852	US-09-352-616A-44	Sequence 44, Appl
26	59.8	4.1	852	US-09-232-149A-44	Sequence 44, Appl
27	58.2	4.0	5427	US-09-009-913-2	Sequence 2, Appl

28	58.2	4.0	5510	US-09-009-913-3	Sequence 3, Appl
29	58.2	4.0	5667	US-09-009-913-4	Sequence 4, Appl
30	55.2	3.8	2975	US-08-368-281-1	Sequence 1, Appl
31	55.2	3.8	3240	US-08-368-281-3	Sequence 3, Appl
32	53.6	3.7	1920	US-08-746-789A-1	Sequence 1, Appl
33	53.4	3.7	1905	US-09-055-113-2	Sequence 2, Appl
34	48.4	3.3	7218	US-08-232-463-14	Sequence 14, Appl
35	47.6	3.3	848	US-09-009-913-338	Sequence 338, App
36	47.4	3.3	7218	US-08-232-463-14	Sequence 14, App
37	38.8	2.7	2280	US-09-009-913-8	Sequence 8, Appl
38	38.8	2.7	2428	US-09-009-913-6	Sequence 6, Appl
39	38.8	2.7	2498	US-09-009-913-10	Sequence 10, Appl
40	35.2	2.4	2432	US-08-469-412A-3	Sequence 3, Appl
41	35.2	2.4	2432	US-09-021-715-3	Sequence 3, Appl
42	33.4	2.3	1212	US-09-046-578-1	Sequence 1, Appl
43	32.6	2.3	1447	US-08-878-177-1	Sequence 1, Appl
44	32.6	2.3	1528	US-08-878-177-3	Sequence 3, Appl
45	32.6	2.3	12886	US-09-453-702B-14	Sequence 14, Appl

#### ALIGNMENTS

RESULT 1					
US-08-878-177-1					
Sequence 1, Application US/08878177					
Patent No. 6294354					
GENERAL INFORMATION:					
APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et					
TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of					
TITLE OF INVENTION: the proteins					
FILE REFERENCE: Chugai Seiyaku Kabushiki Kaisha 5001					
CURRENT FILING DATE: 1997-06-18					
NUMBER OF SEQ ID NOS: 7					
SOFTWARE: PatentIn Ver. 2.0					
SEQ ID NO 1					
LENGTH: 1447					
TYPE: DNA					
ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA					
US-08-878-177-1					
Query Match					
Best Local Similarity 100.0%; Score 1447; DB 4; Length 1447;					
Matches 1447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	GAATTCCGCGAAGATATTTATTTAGCAATTTATTTAGCGATCAATATCTTGATCACA	60		
DB	1	GAATTCCGCGAAGATATTTATTTAGCAATTTATTTAGCGATCAATATCTTGATCACA	60		
QY	61	TTATGCGAAGCACTATTATTAAGAGCAATTTATTTAGTGTGAGTGAAGACCACTTGTGTTG	120		
DB	61	TTATGCGAAGCACTATTATTAAGAGCAATTTATTTAGTGTGAGTGAAGACCACTTGTGTTG	120		
QY	121	AGTGTGCTATGAGTATGCGCCCACTTCAAGAAAGAAATGACGCCCTCTCCAGTG	180		
DB	121	AGTGTGCTATGAGTATGCGCCCACTTCAAGAAAGAAATGACGCCCTCTCCAGTG	180		
QY	181	AATATGCGCAACATCAAGATGAGCCGCGCTTCCCAAGAGATGTTATATCAGC	240		
DB	181	AATATGCGCAACATCAAGATGAGCCGCGCTTCCCAAGAGATGTTATATCAGC	240		
QY	241	CCCCGCGAAGTATCACTTATTAAGTGTGTAACCAACCAAGTTAATGCGTCAAGA	300		
DB	241	CCCCGCGAAGTATCACTTATTAAGTGTGTAACCAACCAAGTTAATGCGTCAAGA	300		
QY	301	ATTCACCTGATGACTGAGGCTGCGCAAGAGGAAATGTTAGAGTTCCAGCAATG	360		
DB	301	ATTCACCTGATGACTGAGGCTGCGCAAGAGGAAATGTTAGAGTTCCAGCAATG	360		
QY	361	TTTGGATGAACTATGAAAGTACATGAAAGAAAGCAATATTCGCGCTCCAAATATGACAA	420		
DB	361	TTTGGATGAACTATGAAAGTACATGAAAGAAAGCAATATTCGCGCTCCAAATATGACAA	420		

OY	421	CCAAATGAACGAAAGGTTTATGTCGACGAGAAATCCAGCTTATGAGGACACACACATGAC	480
Db	421	CCAAATGAACGAAAGGTTTATGTCGACGAGAAATCCAGCTTATGAGGACACACACATGAC	480
OY	481	GGCAGCTGGCTGGAGTGGGCGAGTGAAGAGATATGGTCTTCCAGAGCTGGACATCTGTGTGT	540
Db	481	GGCAGTGGCTGGAGTGGGCGAGTGAAGAGATATGGTCTTCCAGAGCTGGACATCTGTGTGT	540
OY	541	TCGAGAAATATGATGGGAAAGAGTTGTGTAAATATGACCAAGATGACTTCCAGAGACTCA	600
Db	541	TCGAGAAATATGATGGGAAAGAGTTGTGTAAATATGACCAAGATGACTTCCAGAGACTCA	600
OY	601	CGCCGAGCTATTAACGACATATACCTCCCTGCTACACACTTACACTTCCAGAGAGAGAG	660
Db	601	CGCCGAGCTATTAACGACATATACCTCCCTGCTACACACTTACACTTCCAGAGAGAGAG	660
OY	661	CCACTTTTATTTTCCAAATACATCAGTTTACCCAGAAAGCAACGCAAGAAATTAACAACAA	720
Db	661	CCACTTTTATTTTCCAAATACATCAGTTTACCCAGAAAGCAACGCAAGAAATTAACAACAA	720
OY	721	GGCCGAGATTTACCTTATGAGGACGAGAGGAGATACGGCTGGACGAGTGCACAGCCATCCCA	780
Db	721	GGCCGAGATTTACCTTATGAGGACGAGAGGAGATACGGCTGGACGAGTGCACAGCCATCCCA	780
OY	781	CTCAGTCAAAAGCTACCCACATCATCTTCAACAGTGCCTCCAAAAACAGAAAGCAGCGCTC	840
Db	781	CTCAGTCAAAAGCTACCCACATCATCTTCAACAGTGCCTCCAAAAACAGAAAGCAGCGCTC	840
OY	841	CTCAGTTTGAATCCCTTATCAGATTCCTGAGACCGACGAGAGCGCGTTCGCAAAATCCAGGSA	900
Db	841	CTCAGTTTGAATCCCTTATCAGATTCCTGAGACCGACGAGAGCGCGTTCGCAAAATCCAGGSA	900
OY	901	GTGGGAGATACAGCTATGGCAGTTTCTCTACTGAGACTTCTGTGGACAGCTCCAACTCCA	960
Db	901	GTGGGAGATACAGCTATGGCAGTTTCTCTACTGAGACTTCTGTGGACAGCTCCAACTCCA	960
OY	961	ACTGATACCTGGGAGGGACCAAAATGGGAGTTTCAATATCAGACAGACCTCATGAAGTGG	1020
Db	961	ACTGATACCTGGGAGGGACCAAAATGGGAGTTTCAATATCAGACAGACCTCATGAAGTGG	1020
OY	1021	CTCGGCGCTTGGGAGAGAGAAAAAGCAAACTTAACATGAATATATGACAAATCCAGCGCTG	1080
Db	1021	CTCGGCGCTTGGGAGAGAGAAAAAGCAAACTTAACATGAATATATGACAAATCCAGCGCTG	1080
OY	1081	CACCTTCGCTACTACTATGACAAATAATATATGACTAAAGTTTCATGTAACGCTATGCGCT	1140
Db	1081	CACCTTCGCTACTACTATGACAAATAATATATGACTAAAGTTTCATGTAACGCTATGCGCT	1140
OY	1141	ACAAATTTGATTTCCAGGAGATGCGCTCAGGCGCTCCAGCGCTCACCGCTCCACAATCATCCA	1200
Db	1141	ACAAATTTGATTTCCAGGAGATGCGCTCAGGCGCTCCAGCGCTCACCGCTCCACAATCATCCA	1200
OY	1201	TGTACAAATATACCACATCAGACCTCCCTCAATCAGATGATCTTACATGACACACCCAGAGAA	1260
Db	1201	TGTACAAATATACCACATCAGACCTCCCTCAATCAGATGATCTTACATGACACACCCAGAGAA	1260
OY	1261	TGAAGTTTGTATGCTCCCCATCCCGCTGTTGGCCCGTAACCTCATCCAGCTTTTGTGCTG	1320
Db	1261	TGAAGTTTGTATGCTCCCCATCCCGCTGTTGGCCCGTAACCTCATCCAGCTTTTGTGCTG	1320
OY	1321	CCCCTAATTCATCTGAGATTCACCAACTGAGGAGCATCTACCCCAATACAGGCGTGCAG	1380
Db	1321	CCCCTAATTCATCTGAGATTCACCAACTGAGGAGCATCTACCCCAATACAGGCGTGCAG	1380
OY	1381	CTGCTCATATGCTTCCCATCTTGTGGACCTTACTAATGATGGGAGAAAGAAAGAAAGCGCC	1440
Db	1381	CTGCTCATATGCTTCCCATCTTGTGGACCTTACTAATGATGGGAGAAAGAAAGAAAGCGCC	1440
OY	1441	AAAAAA 1447	
Db	1441	AAAAAA 1447	

```

RESULT 2
US-08-878-177-3
; Sequence 3, Application US/08878177
; Patent No. 6294354
; GENERAL INFORMATION:
; APPLICANT: Chugai Seliyaku Kabushiki Kaisha, 5001, Yamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; TITLE OF INVENTION: the Proteins
; FILE REFERENCE: chugai seliyaku kabushiki kaisha 5001
; CURRENT APPLICATION NUMBER: US/08/878,177
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: C-eryg gene, chicken DNA
US-08-878-177-3

```

Query Match	Similarity	93.7%	Score 1356	DB 4	Length 1528
Best Local Match	Similarity	94.7%	Pred. No. 0		
Matches 1447	Conservative	0	Mismatches	0	Indels 81
					Gaps 1
QY	1	GAATTCGCGCAACGATAATATTATTTATAGCAATTTATAGCATCATATATCTTGATGACA	60		
Db	1	GAATTCGCGCAACGATAATATTATTTATAGCAATTTATAGCATCATATATCTTGATGACA	60		
QY	61	TTATGGCAAGACATTTTAAGGAGCATTTATCAGTGTGATGTAAGAACACAGTCTGTGG	120		
Db	61	TTATGGCAAGACATTTTAAGGAGCATTTATCAGTGTGATGTAAGAACACAGTCTGTGG	120		
QY	121	AGTGTGCTTACGAGATCGCCCACTTGCAGAAAGACGAATATACAGCTCTCTTCAGTGTG	180		
Db	121	AGTGTGCTTACGAGATCGCCCACTTGCAGAAAGACGAATATACAGCTCTCTTCAGTGTG	180		
QY	181	AATATGGCAAAACATCAAAAGATGAGCCGCGCTTCCACAGACATGTGTTATCACAGC	240		
Db	181	AATATGGCAAAACATCAAAAGATGAGCCGCGCTTCCACAGACATGTGTTATCACAGC	240		
QY	241	CCCCGGCCAGATTTACCATTTAAGATGAGGTGTAAACCCAAACAGTTAATGGGTCAAGCA	300		
Db	241	CCCCGGCCAGATTTACCATTTAAGATGAGGTGTAAACCCAAACAGTTAATGGGTCAAGCA	300		
QY	301	ATTTCACCTGATGACACTGCAGCGCTGGCGAAAAGAGGGAAAATGTTTATAGACTTCAGACATG	360		
Db	301	ATTTCACCTGATGACACTGCAGCGCTGGCGAAAAGAGGGAAAATGTTTATAGACTTCAGACATG	360		
QY	361	TTGGGATGACATATGGAAGCTACATGTGAAGAGAAACATATTCGCGCTCCAAATATGACAA	420		
Db	361	TTGGGATGACATATGGAAGCTACATGTGAAGAGAAACATATTCGCGCTCCAAATATGACAA	420		
QY	421	CCATGAAACGAAGAGTTTGTGGCAGAGATCCTACGTTATGGAGCACACACCATGTATC	480		
Db	421	CCATGAAACGAAGAGTTTGTGGCAGAGATCCTACGTTATGGAGCACACACCATGTATC	480		
QY	481	GGCAGTGGCTTGAATGGGCACTGGAAGAGTATGGTCTTCCAGACGTGACATCTCTGTGT	540		
Db	481	GGCAGTGGCTTGAATGGGCACTGGAAGAGTATGGTCTTCCAGACGTGACATCTCTGTGT	540		
QY	541	TCAGAACATTTGATGGGAAAAGATTGTGTAAATATGACCAAAAGATGACTTCAGAGACATCA	600		
Db	541	TCAGAACATTTGATGGGAAAAGATTGTGTAAATATGACCAAAAGATGACTTCAGAGACATCA	600		
QY	601	CGCGAGCTATTAACGACAGATATCTCTCTGTACACACTTACACTACTCTCAGAGAGA-----	654		
Db	601	CGCGAGCTATTAACGACAGATATCTCTCTGTACACACTTACACTACTCTCAGAGAGACTCTC	660		
QY	655	-----GAGGAGCCCACTTTATTTATTTTTCCAAAATACATCACTTATACCGAGAG	699		
Db	661	TTCCACATTTGACTTTCAGATGATGTGTATTAAGGCGCTTACAAACCTCTCCACGGTTAATGC	720		
QY	655	-----GAGGAGCCCACTTTATTTATTTTTCCAAAATACATCACTTATACCGAGAG	699		

```

Db 721 ATGTAGAAACAGAGAGAGACCACTTTATTTTCCAAATATACAGTTTACCCAGAG 780
Qy 700 CAAGCAAAAGATTAACAACAAGGCCAGATTACCTTATGAGCAAGCGAGATGAGCT 759
Db 781 CACGCAAAAGATTAACAACAAGGCCAGATTATTTAGACCAAGCGAGATGAGCT 840
Qy 760 GGAGAGATCACAGCATCCCTCAGTCAAAAGCTACCAACCATCTTCAACAGTGC 819
Db 841 GGAGAGATCACAGCATCCCTCAGTCAAAAGCTACCAACCATCTTCAACAGTGC 900
Qy 820 CCAAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879
Db 901 CCAAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 880 GCGCTCTTGCAATTCAGAGAGAGAGAGAGAGATACAGTATGAGCAGTCTTCTAGAGCTTC 939
Db 961 GCGCTCTTGCAATTCAGAGAGAGAGAGAGAGATACAGTATGAGCAGTCTTCTAGAGCTTC 1020
Qy 940 TTGTGAGAGAGCTCAACTCAACTGACATGACCTGGAGAGAGAGAGAGAGAGAGAGAG 999
Db 1021 TTGTGAGAGAGCTCAACTCAACTGACATGACCTGGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1000 TGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1059
Db 1081 TGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Qy 1060 ACTATGACAAATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1119
Db 1141 ACTATGACAAATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Qy 1120 TTCAATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179
Db 1201 TTCAATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy 1180 CTCACCTCCAGATTCATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1239
Db 1261 CTCACCTCCAGATTCATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1240 ACCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1299
Db 1321 ACCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1300 CCTATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1359
Db 1381 CCTATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy 1360 ACCCAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1419
Db 1441 ACCCAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy 1420 TGGGGAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1447
Db 1501 TGGGGAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1528

```

```

RESULT 3
US-08-343-443B-3
Sequence 3, Application US/08343443B
Patent No. 5968734
GENERAL INFORMATION:
APPLICANT: Aurias, Alain
APPLICANT: Delattre, Olivier
APPLICANT: Desmaze, Chantal
APPLICANT: Melot, Thomas
APPLICANT: Ploougas, Beatrice
APPLICANT: Thomas, Gilles
APPLICANT: Zucman, Jessica
TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
TITLE OF INVENTION: TRANSLOCATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID

```

```

TITLE OF INVENTION: TRANSLOCATIONS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates
STREET: 230 South Fifteenth Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: AEDIT 1.0 DOS text editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,443B
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00494
FILING DATE: 19-MAY-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: FR 92/06123
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989,6121P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2938 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 143..1498
US-08-343-443B-3
Query Match
Best Local Similarity 66.5%: Score 563.8; DB 2: Length 2938;
Matches 913; Conservative 0; Mismatches 427; Indels 33; Gaps 6;
Qy 63 ATGCAAGCACTATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122
Db 143 ATGCAAGCACTATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 202
Qy 123 TGTGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179
Db 203 TCAACGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 262
Qy 180 GAATATGAGCAAAACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 239
Db 263 GACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 322
Qy 240 CCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 299
Db 323 CC-----AGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373
Qy 300 AATTACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
Db 374 GAGTCTCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
Qy 360 GTTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416
Db 434 AACCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 493
Qy 417 ACAACCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476
Db 494 ACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 553

```

```

QY 477 GTACGAGTGGCTGGAGTGGGAGTGAAGAGTATGCTCTCCAGACGTGGACATCTTG 536
D 554 GTGAGGCAATGGCTGGAGTGGGAGTGAAGAGTATGCTTGATGAGAGATGACATATCC 613
QY 537 TTGTTCAGAACTGATGGGAAAGAGTGTGTGTAATGACCAAGATGACTTCCAGAGA 596
D 614 TTTTTCAGAACTGATGGGAAAGAGTGTGTGTAATGAAAGAGGAGGAGTCTTCCGCG 673
QY 597 CTCAGCGGAGCTATTAACGAGATATCTCTGTCACACCTACACTCTCAGAGAGAGA 656
D 674 GCCACACCTCTCTACACGAGAGTGTGTGTACACCTCAGTCTACAGGAGAG- 732
QY 657 GGAGCCACTTTATTTTCCAAATACATCATGTTTACCAGAGAACGCAAGCAATACA 716
D 733 -----TTCACTGTGGCTATTAATACAACTCCACACCGCACCAATCTCAGATGAGT 787
QY 717 ACAAGGCCATTTACTTATGAGCAAGCAGAGATCAGCGTGGACGATCAGACCAT 776
D 788 GTCAAGAGAACCTTTCTTATGACTCAGAGAGAGAGAGAGAGAGCAATTAACATGAT 847
QY 777 CC---CACTCAGTCAAAAGCTACCCACATCATCTTCAACAGTGGCCAAACAGAGAC 833
D 848 TCTGGCTCAACAAATCTCTCCCTTGAGGGGCAAAACGATCAGTAAGATACAGAG 907
QY 834 CAGCGTCTCAAGTATGATCTTATCAGATTTCTTGACCGACAGCAGCCGCTTGCAAT 893
D 908 CAACGGCCGAGCAGATCCGTATCAGATCTGGGGCCGACAGCAGATCGCTAGCCAC 967
QY 894 CAGAGGAGTGGGAGATACAGATGAGTCTGCTACTGAGAGCTTCTGTGGAGAGCTCC 953
D 968 CCTGGAAGCGGAGATCAGATGAGTGGCAATCTCTGAGAGTCTCTCGAGAGGCGC 1027
QY 954 AACTCAACTGATCACTCGGAGAGGCAAAATGGGAGTTCAGAGATGACAGACCTGAT 1013
D 1028 AACCGCAGCTGTATCACTGGAGGAGGACCAAGGAGGATTCAAATGAGAGGACCCGAT 1087
QY 1014 GAATGGCTGGCGTGGGAGAGAGAGAAAGCAAACTTAACATGAATGACAAATC 1073
D 1088 GAGGTGGCCAGGGCTGGGGCGAGCGGAAAGCAAGCCAAATGATTAAGACAACTG 1147
QY 1074 AGCGTGCATCTGCTCTACTATGACAAATATTTGATGAATGATGATGATGAAACG 1133
D 1148 AGCGGGGCTCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1207
QY 1134 TATGCTCAAAATTTGATTTCCAGAGATGCTCAGGCTCAGGCTCAGGCTCAGGAA 1193
D 1208 TATGCTCAAAATTTGATTTCCAGAGATGCTCAGGCTCAGGCTCAGGCTCAGGAA 1267
QY 1194 TCATCCATGTAACAAATACCATGACATGCTCCCTCAATGAGTTCATGACACACCC 1253
D 1268 TCGTCCATGTAACAAATACCATGACATGCTCCCTCAATGAGTTCATGACACACCC 1327
QY 1254 CAGAAGATGAATTTGAGTTCGCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313
D 1328 CAGAAGATGAATTTGAGTTCGCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1387
QY 1314 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1371
D 1388 TTTGAGGCGCATACATGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1447
QY 1372 -----GGCTGCGAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1417
D 1448 GTCCCCCGCATCTTAACACCCAGCTGCTTACACTTACAGCTGCTGCTGCTGCTGCT 1500

```

```

: APPLICANT: Hendricks, Timothy J.
: TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
: FILE REFERENCE: CASE-03828
: CURRENT APPLICATION NUMBER: US/09/360,779
: CURRENT FILING DATE: 1999-07-26
: EARLIER APPLICATION NUMBER: 60/094,264
: EARLIER FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1752
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (112)..(1131)
US-09-360-779-1
Query Match 12.4%; Score 180; DB 4; Length 1752;
Best Local Similarity 74.0%; Pred. No. 1.5e-51;
Matches 228; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 879 AGCGTCTGCAAAATCCAGGAGTGGGAGATACAGTATGCGAGTCTTCTGAGCTT 938
D 529 AGCCCTCGGTACAGAAAGCAGCGGAGATCCAGTGTGCGAGTCTTCTGAGCTT 588
QY 939 CTGTGGGAGCTCCACCTCAGTCACTGATCCAGTGGGAGGAGCAAAATGGGAGTTCAG 998
D 589 CTGGCAGACCGCGGAGAGCGCGGTGCTGATGCGTGGGAGGAGGCGCCAGGAGTTCAG 648
QY 999 ATGACAGACCTGATGAGTGGCTGCGCTGGGAGAGAGAGAAAGCAAACTTAACATG 1058
D 649 CTACAGCAGCCCGAGAGGAGTGGGCGAGCGTGGGAGGAGGAGCAAGCAAGCCAAATG 708
QY 1059 AACTATGACAACTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
D 709 AACTATGACAACTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768
QY 1119 GTTCATGTAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1178
D 769 GTGACAGGAGAGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 828
QY 1179 CCTGACCC 1186
D 829 CCACACCC 836

```

```

RESULT 4
US-09-360-779-1
: Sequence 1, Application US/09360779
: Patent No. 6268216
: GENERAL INFORMATION:
: APPLICANT: Demerits, Evan S.
: APPLICANT: Fyodorov, Dmitry V.

```

```

RESULT 5
US-09-435-335-1
: Sequence 1, Application US/09435335
: Patent No. 6384204
: GENERAL INFORMATION:
: APPLICANT: Demerits, Evan S.
: APPLICANT: Fyodorov, Dmitry V.
: APPLICANT: Hendricks, Timothy J.
: TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
: FILE REFERENCE: CASE-04027
: CURRENT APPLICATION NUMBER: US/09/435,335
: CURRENT FILING DATE: 1999-11-05
: EARLIER APPLICATION NUMBER: 09/360,779
: EARLIER FILING DATE: 1999-07-26
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1752
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (112)..(1131)
US-09-435-335-1

```





	Best Local Similarity	61.6%	Pred. No. 2e-32;		Matches 199;	Conservative 0;	Mismatches 124;	Indels 0;	Gaps 0;
QY	835	AGCGTCTCACTTATGATTCCTTATAGATTTCTTGGACCGACGACGACGCCGTTGGCAATTC	894						
Db	1207	ACCGTGTGACCTCAATTAAGGACAAGCCTGTCAATTCCTGCTGCTGCCCTAGCTGGCTACCA	1266						
QY	895	CAGGAGATGGGCGACATATACAGTTATGGCAGTTCTCTACTAGAGCTTCTGTGGACAGCTCCA	954						
Db	1267	CAGGCAATGGAGCAATTCACGCTATATGGCAGTTCTCTTGAGATTACTACTATATAATCT	1326						
QY	955	ACTCCAACTGCATCACTCGTGGGAGGGCACAAATGGGAGTTCAGATGACAGACCTGATG	1014						
Db	1327	GTCAGTCTTTTATATAGCTGTGGACAGAGATGGCTGGGAATTCCTCACTCCACAGATG	1386						
QY	1015	AAGTGGCTCGGCGCTGGGGGAGAGAGAAACCAAACTCATGATGACATGACAAATCCA	1074						
Db	1387	AGGTGGCCAGAGATGGGGAAGAGGAAAACCAAACTTAAGATGATATTATAGAAATCGA	1446						
QY	1075	GCCGTGCACCTTGGCTACTACTATGACAAAAATATTATGACTAAAGTTCATGGTAAACGCT	1134						
Db	1447	GCCGTGGCGCTACGCTACTATTACGACAAAAAATCATTCACAAAGCAGCGGGGAAACGCT	1506						
QY	1135	ATGCTACAAATTTGATTTCCAC	1157						
Db	1507	ACGTGTACCGCTTGTGTGTGAC	1529						

```

RESULT 8
US-09-344-579-1
: Sequence 1, Application US/09344579
: Patent No. 6054316
:
: GENERAL INFORMATION:
: APPLICANT: Brenda F. Baker
: APPLICANT: Lex M. Cowsett
: TITLE OF INVENTION: ANTISENSE MODULATION OF ETS-2 EXPRESSION
: FILE REFERENCE: RTS-0063
: CURRENT APPLICATION NUMBER: US/09/344,579
: CURRENT FILING DATE: 1999-06-25
: NUMBER OF SEQ ID NOS: 47
: SEQ ID NO 1
:
: LENGTH: 2268
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (291)..(1700)
: US-09-344-579-1

```

	Query Match Best Local Similarity	8.4% 66.5%	Score 122.2; Pred. No. 1.7e-31;	DB 3;	Length 2266;
	Matches	Conservative	0;	Mismatches 88;	Indels 0;
					Gaps 0;
QY	895	CAGGAGCTGGGCAATACAGCTATATGGCAGTTCCTACTGTGAGCTTCTGTGGACAGCTCCA			
Db	1363	CAGGAAGTGGACCTATTTCAGCTGTGGCAGTTTCTCTCTGTGAGCTGCTATTACAGCAAAATCCT			
QY	955	ACTCCAACTGCATCACCCTGGGAGGGGCAAAATGGGGAGTTCAAGATGACAGACCCCTGATG			
Db	1423	GCCAGTCATTCATCAGCTGAGCTGGAGACGGATGGGAGTTTAAAGTCTGCCGACCCCGATG			
QY	1015	AAGTGGCTCGGCGCTTTGGGGAGAGAGAAAGCAAACTTAACATGTAATATGACAAATCTCA			
Db	1483	AGGTGGCCCGCGCGTGGGGAAAGAGGAAAAATTAAGCCCAAGATGAACTACAGAGAAAGCTGA			
QY	1075	GCCTGTGCACCTTCGCTACTACTACTATACCAAAATATTATGACTTAAGTTATGTATAACGCT			
Db	1543	GCCTGGGCGCTTAAGCTACTATTATACCAAGAAACATATCATCCACAGACGCTGGGGGAAGCGCT			
QY	1135	ATGCCCTACAAATTGATTTCCAC			
Db	1603	ACGTGTACCGCTTCGTGTGGAC			

```

1      RESULT 9
2      US-08-469-412A-1
3      Sequence 1, Application US/08469412A
4      Patent No. 5856125
5      GENERAL INFORMATION:
6      APPLICANT: Mavrothallastis, George J.
7      APPLICANT: Blair, Donald G.
8      APPLICANT: Fisher, Robert J.
9      APPLICANT: Beal Jr., Gregory J.
10     APPLICANT: Athanasiou, Meroyl A.
11     APPLICANT: Sgouras, Dionysios N.
12     TITLE OF INVENTION: The ERF Genetic Locus and Its Products
13     NUMBER OF SEQUENCES: 16
14     CORRESPONDENCE ADDRESS:
15     ADDRESSEE: Townsend and Townsend and Crew LLP
16     STREET: Two Embarcadero Center, Eighth Floor
17     CITY: San Francisco
18     STATE: California
19     COUNTRY: USA
20     ZIP: 94111-3834
21     COMPUTER READABLE FORM:
22     MEDIUM TYPE: Floppy disk
23     COMPUTER: IBM PC compatible
24     OPERATING SYSTEM: PC-DOS/MS-DOS
25     SOFTWARE: Patentln Release #1.0, Version #1.30
26     CURRENT APPLICATION DATA:
27     APPLICATION NUMBER: US/08/469,412A
28     FILING DATE: 05-JUN-1995
29     CLASSIFICATION: 435
30     ATTORNEY/AGENT INFORMATION:
31     NAME: Garrett-Wackowski, Eugenia
32     REGISTRATION NUMBER: 37,330
33     REFERENCE/DOCKET NUMBER: 015280-229000
34     TELECOMMUNICATION INFORMATION:
35     TELEPHONE: (415) 576-0200
36     TELEFAX: (415) 576-0300
37     INFORMATION FOR SEQ ID NO: 1:
38     SEQUENCE CHARACTERISTICS:
39     LENGTH: 2667 base pairs
40     TYPE: nucleic acid
41     STRANDEDNESS: single
42     TOPOLOGY: linear
43     MOLECULE TYPE: CDNA
44     FEATURE:
45     NAME/KEY: CDS
46     LOCATION: 123..1769
47     OTHER INFORMATION: /note="human ERF (ETS2 Repressor
48     OTHER INFORMATION: Factor) CDNA"
49     US-08-469-412A-1

```

	Query Match	8.1%;	Score 117.4;	DB 2;	Length 2667;
	Best Local Similarity	63.1%;	Pred. No. 8.8e-30;		
	Matches	181;	Conservative	0;	Mismatches 106; Indels 0; Gaps 0;
QY	894	CCAGGAGATGGGACAGATACAGCTATGGCAGTTCCTACTGAGAGCTTCTGCGACAGCACC	953		
Db	186	CCGTGGCTCAAGAGCAGATACAGCTGTGGACCTTTATCTGTGGAGCTGCGGAAGGAGG	245		
QY	954	AATCTCAACTGCTATCACCCTGGGAGGGCCAAATGGGAGTTCAGATGACAGACCTTCAT	1013		
Db	246	TACAGGGGCGCTATGGCTGTGGCAGGGGAGTACGAGGGGAATTCGTATCAAGAGACCTTCAT	305		
QY	1014	GAACTGGCTGGGGCTTGGGGGAGACAGGAAAGCAAACTTAATCATGAACTATGACAAACTC	1073		
Db	306	GAGGTGGCCCGGCGTGTGGGGCGTTCGCAAGTGCAGAGCCCGAGATCAATTACGACAAAGCTG	365		
QY	1074	AGCGTGCACTTCCCTACTACTATGACAAAAAATATTATGACATAAGTTCATGTGTAACGC	1133		
Db	366	AGCGGGGCGCTGCCTATTACTATTACAGGCGCATTCGCACAAAGGCAAGGGGAAACGG	425		
QY	1134	TATGCGTACAAATTTGATTTCCACAGGAATCGCTCAGCGCCCTCCAGCC	1180		



QY 790 AAGTACCCACATCTTCAACAGTGGCCCAAGAGACGAGCTCTCATGTTAG 849  
Db 125 GTCCCTCCCTTGGAGGGGACAAAGATAGTAAGTACAGAGCAAGGCCAGCCAG 184  
QY 850 ATCTTAATGAGATCTTGGAGCCGACGAGCCGCTTGCAGATCCAGGGAGTGGGAGA 909  
Db 185 ATCCGTATGACATCTTGGAGCCGACGAGCCGCTTGCAGATCCAGGGAGTGGGAGA 244  
QY 910 TACAGCTATGAGCTTCTTACTGAGCTTCTGTCGAGACGCTCCACTCCATGATCA 969  
Db 245 TCCAGCTGTGGCATTTCTCTCTGAGCTGCTCTCCGAGGAGCCAGCCAGCTGTATCA 304  
QY 970 CCTGGAGGCGACAAATGGGAGT 993  
Db 305 CCTGGAGGCGACCAAGCGGGAGT 328

## RESULT 12

US-08-875-944B-1  
Sequence 1, Application US/08875944B  
Patent No. 6096542  
GENERAL INFORMATION:  
APPLICANT: FUJINAGA, Kei  
APPLICANT: YOSHIDA, Koichi  
APPLICANT: HIGASHINO, Fumihito  
TITLE OF INVENTION: CANCER CONTROL  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
STREET: 624 Ninth Street N.W., Ste. 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,944B  
FILING DATE: 07-AUG-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 07-020173  
FILING DATE: 08-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/00016  
FILING DATE: 09-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: FUJINAGA-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-5197  
TELEFAX: (202) 737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2064 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1386  
US-08-875-944B-1

Query Match 6.6%; Score 96; DB 3; Length 2064;  
Best Local Similarity 62.5%; Pred. No. 1.8e-22;  
Matches 150; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
QY 910 TACAGCTATGAGCTTCTTACTGAGCTTCTGTCGAGACGCTCCACTCCATGATCA 969

Db 956 TGCAGCTGTGGCAATTTCTGTGGCTTGTGAGATGACCCAAAGATGCCATTTCAATG 1015  
QY 970 CCTGGAGGCGCAAAATGGGAGTTTCAAGATGACAGACCCCTGATGAGTGGCTGGCGGT 1029  
Db 1016 CTTGGAGGCGCCGGGGAATGAGTTTCAAGCTTCACTTGAAGCTTGAGAGAGTCCGAGCTCT 1075  
QY 1030 GGGGAGAGAGGAAAGCAAACTTAACATGAACTTAACAACTGAGCGCTGCACTTGCT 1089  
Db 1076 GGGGCAATCCAAAGAACCGCGCAGCATGAAATGAGAACTGAGCGCGCTGCCGAT 1135  
QY 1090 ACTACTATGACAAAATTTATGACTTAAGTTGATGCTTAAGCTTAAGCTTAAGCTTAAGTTG 1149  
Db 1136 ACTATTATGAGAAAGGATCATGACGAAGGTGGCTGTGAGCGCTTACGTTGACAAATTTG 1195

## RESULT 13

US-09-116-049-3  
Sequence 3, Application US/09116049A  
Patent No. 6248351  
GENERAL INFORMATION:  
APPLICANT: Hung, Men-Chie  
TITLE OF INVENTION: HUMAN PEA3 IS A TUMOR SUPPRESSOR FOR CANCER CELLS  
FILE REFERENCE: UTSC:582  
CURRENT APPLICATION NUMBER: US/09/116,049A  
CURRENT FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2064  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-116-049-3

Query Match 6.6%; Score 96; DB 4; Length 2064;  
Best Local Similarity 62.5%; Pred. No. 1.8e-22;  
Matches 150; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 910 TACAGCTATGAGCTTCTTACTGAGCTTGTGTCGAGACGCTCCACTCCATGATCA 969  
Db 956 TGCAGCTGTGGCAATTTCTGTGGCTTGTGAGATGACCCAAAGATGCCATTTCAATG 1015  
QY 970 CCTGGAGGCGCAAAATGGGAGTTTCAAGATGACAGACCTGATGAAGTGGCTGGCGGT 1029  
Db 1016 CTTGGAGGCGCCGGGGAATGAGTTTCAAGCTTCACTTGAAGCTTGAGAGAGTCCGAGCTCT 1075  
QY 1030 GGGGAGAGAGGAAAGCAAACTTAACATGAACTTAACAACTGAGCGCTGCACTTGCT 1089  
Db 1076 GGGGCAATCCAAAGAACCGCGCAGCATGAAATGAGAACTGAGCGCGCTGCCGAT 1135  
QY 1090 ACTACTATGACAAAATTTATGACTTAAGTTGATGCTTAAGCTTAAGCTTAAGCTTAAGTTG 1149  
Db 1136 ACTATTATGAGAAAGGATCATGACGAAGGTGGCTGTGAGCGCTTACGTTGACAAATTTG 1195

## RESULT 14

US-08-780-835B-1  
Sequence 1, Application US/08780835B  
Patent No. 5922688  
GENERAL INFORMATION:  
APPLICANT: Hung, Men-Chie  
APPLICANT: Xing, Xiangming  
TITLE OF INVENTION: PEA3 is a Tumor Suppressor  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE AND DURKEE  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210-4433

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 15-  
US-09-303-268-1  
Sequence 1, Application US/09303268  
Patent No. 6172212  
GENERAL INFORMATION:  
APPLICANT: Hung, Mien-Chie  
Xing, Xiangming  
TITLE OF INVENTION: PEAK Is a Tumor Suppressor  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE AND DUKKEE  
STREET: P.O. BOX 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/303,268  
FILING DATE: 30-Apr-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

```

QY      865  TTGGACCGACGACGACCGCTTTGCAATTCACGAGGAGGCGACAGTACAGGTATGGCAGT  924
Db      1315  TTGGAGAGCTTTCGGGGAGGGGGCCACCCTTACGACGGCCCGGGGTGTGCTTACAACTGTGGCAGT  1374

QY      925  TCTCTACTGAGAGCTTCTGTGCGACAGCTTCCAACTCCAACTGCATCACTCGGAGGGACACA  984
Db      1375  TTTCGTGGTGGCCCTCTGTGATGACCCCAACAAATGTGTCATTTGATTTGCTTGGACAGGCCGG  1433

QY      985  ATGGGAGATTCAAGATGACAGACCCGTGTAAGTGGCTCGGCGTGTGGGGAAGAGAGAAA  1044
Db      1435  GAATGGAGATTAAACTAATTGTAACCTGAAGAAGTTGCCAGGCTGTGGGGTTATCCAGAA  1494

QY      1045  GCAAACTTACATGACATATGACAAACTCAGCCGTGCACCTTGGCTACTACTATGCAAAA  1104
Db      1495  ACCGGCCACGACATTAATTATGACAAAGCTGACGCCCTGCTCGATACATTAATGGAAG  1554

QY      1105  ATATTATGACATAAAGTTCATGTTAAAGCCTATGGCTATCAAAATTTGATTTCGA  1156
Db      1555  GCATCATGACAAAGAGTGGCTGGCGAAGCGTACGTTGTAACAATTTGTGTGCGA  1606

Search completed: July 28, 2003, 12:16:34
Job time : 93.3059 secs

```

Search completed: July 28, 2003, 12:16:34  
Job time : 93.3059 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 10:09:50 ; Search time 398.967 Seconds  
(without alignments)  
7482.227 Million cell updates/sec

Title: US-09-902-772-1

Perfect score: 1447  
Sequence: 1 gaattccgcgaacgaataat.....gaaagaagcgaagaataa 1447

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCITUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1447	100.0	1447	11	US-09-902-772-1
2	1356	93.7	1528	11	US-09-902-772-3
3	921	63.6	3166	15	US-10-205-823-98
4	387	26.7	567	10	US-09-864-761-18410
5	354.6	24.5	473	10	US-09-864-761-20472
6	307	21.2	467	12	US-09-918-995-23356
7	276.6	19.1	549	10	US-09-923-779-54
8	255	17.6	420	10	US-09-864-761-3705
9	255	17.6	454	10	US-09-864-761-1653
10	180	12.4	1752	10	US-09-850-799-1
11	167.8	11.6	553	10	US-09-864-761-7427
12	163.2	11.3	348	10	US-09-864-761-29364
13	163.2	11.3	477	10	US-09-864-761-5035
14	163.2	11.3	519	10	US-09-864-761-12798
15	162.2	11.2	225	10	US-09-864-761-21761
16	148.4	10.3	2532	15	US-10-037-270-869

17	122.8	8.5	416	12	US-09-918-995-8472	Sequence 8472, Ap
18	122.4	8.5	165	10	US-09-864-761-23294	Sequence 23294, A
19	122.2	8.4	473	12	US-09-918-995-27904	Sequence 27904, A
20	122.2	8.4	1884	11	US-09-925-300-420	Sequence 420, App
21	122.2	8.4	2188	15	US-09-920-300A-1716	Sequence 1716, Ap
22	122.2	8.4	2188	15	US-10-033-528-1716	Sequence 1716, Ap
23	122.2	8.4	2268	11	US-09-920-300A-1693	Sequence 1693, Ap
24	122.2	8.4	2268	15	US-10-033-528-1693	Sequence 1693, Ap
25	122.2	8.4	2269	11	US-09-954-531-955	Sequence 955, App
26	122.2	8.4	3692	15	US-10-106-698-824	Sequence 824, App
27	115.8	8.0	3178	15	US-10-108-605-124	Sequence 124, App
28	115.8	8.0	3178	15	US-10-108-605-128	Sequence 128, App
29	113	7.8	2952	10	US-09-925-297-257	Sequence 257, App
30	113	7.8	4106	15	US-10-084-817-187	Sequence 187, App
31	101	7.0	300	15	US-10-177-063-15	Sequence 15, App1
32	99.4	6.9	300	15	US-10-177-063-16	Sequence 16, App1
33	96	6.6	2064	12	US-09-884-363-3	Sequence 3, App11
34	96	6.6	2064	15	US-09-884-363-1	Sequence 1788, Ap
35	96	6.6	2333	11	US-09-920-300A-1788	Sequence 3316, Ap
36	96	6.6	2333	11	US-09-880-107-3316	Sequence 1788, Ap
37	96	6.6	2333	15	US-10-033-528-1788	Sequence 1788, Ap
38	95.2	6.6	2410	12	US-09-884-363-1	Sequence 1788, Ap
39	90.6	6.3	2256	15	US-10-098-841-317	Sequence 317, App
40	75.6	5.2	180	15	US-10-177-063-27	Sequence 27, App1
41	70.4	4.9	4190	11	US-09-962-832-111	Sequence 111, App
42	67.6	4.7	65042	15	US-10-329-124-3	Sequence 3, App11
43	66.2	4.6	414	10	US-09-864-761-10426	Sequence 10426, A
44	65.8	4.5	303	10	US-09-864-761-27055	Sequence 27055, A
45	63.2	4.4	496	12	US-09-918-995-6907	Sequence 6907, Ap

#### ALIGNMENTS

RESULT 1  
US-09-902-772-1  
Sequence 1, Application us/09902772  
Patent No. US20020164739A1  
GENERAL INFORMATION:  
APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et  
TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
FILE REFERENCE: chugai selyaku kabushiki kaisha 5001  
CURRENT APPLICATION NUMBER: US/09/902,772  
CURRENT FILING DATE: 2001-07-12  
PRIOR APPLICATION NUMBER: US/08/878,177  
PRIOR FILING DATE: 1997-06-18  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 1447  
TYPE: DNA  
ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA  
US-09-902-772-1

Query Match	100.0%	Score 1447;	DB 11;	Length 1447;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1447;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	GAATTCGCGAAGCAATATATATATAGCAATATATAGGCAATATATCTTGATACA	60		
1	GAATTCGCGAAGCAATATATATATATAGCAATATATAGGCAATATATCTTGATACA	60		
1	GAATTCGCGAAGCAATATATATATATAGCAATATATAGGCAATATATCTTGATACA	60		
61	TTATGCGAAGCAATATATATATAGCAATATATAGGCAATATATCTTGATACA	120		
61	TTATGCGAAGCAATATATATATAGCAATATATAGGCAATATATCTTGATACA	120		
61	TTATGCGAAGCAATATATATATAGCAATATATAGGCAATATATCTTGATACA	120		
121	AGTGTGCTTACGATGCGCCCTTCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	180		
121	AGTGTGCTTACGATGCGCCCTTCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	180		
121	AGTGTGCTTACGATGCGCCCTTCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	180		
181	AATATGCGAAGCAATATATATATAGCAATATATAGGCAATATATCTTGATACA	240		

Dp	181	AATATGGGCAAAACATCAAAAGTGAAGCCGGCGTTCCCCAGACGACTGGTATATACAGC	240
OY	241	CCCCGGCCAGAGTTACCACTTAAGATGAGTGTAAACCAACAGGTTAATGGGTCAAGGA	300
Dp	241	CCCCGGCCAGAGTTACCACTTAAGATGAGTGTAAACCAACAGGTTAATGGGTCAAGGA	300
OY	301	ATTCAACCGATGACACGCGAGCGGTGGCAAAAGGAGGGAATAAGTTATGCAAGTTACAGCATTG	360
OY	361	TTGGGATGAACATGATGGAAGCTACATATGGAAGAGAAACATATTCGCCCTCCAAATATGACAA	420
Dp	361	TTGGGATGAACATGATGGAAGCTACATATGGAAGAGAAACATATTCGCCCTCCAAATATGACAA	420
OY	421	CCAAATGAAGAGAGAGTTATGTGTGCCAGAGATCCTACGTTATGAGACACAGACCATGTATC	480
Dp	421	CCAAATGAAGAGAGAGTTATGTGTGCCAGAGATCCTACGTTATGAGACACAGACCATGTATC	480
OY	481	GGCAGTGGCTGGAGTGGGCGACTGGAAGAGATGTGGTCTTCAGAGCTGGACATCTTGTGT	540
Dp	481	GGCAGTGGCTGGAGTGGGCGACTGGAAGAGATGTGGTCTTCAGAGCTGGACATCTTGTGT	540
OY	541	TCCGAACATTTGATGGGAAAGAGTGTGTAAATATACCAAGATATCTTCCAGAGACTCA	600
Dp	541	TCCGAACATTTGATGGGAAAGAGTGTGTAAATATACCAAGATATCTTCCAGAGACTCA	600
OY	601	CGCCGAGCTATTAACGACAGATATCCTCTGTACACCTACACTACCTCAGAGAGAGAGAG	660
Dp	601	CGCCGAGCTATTAACGACAGATATCCTCTGTACACCTACACTACCTCAGAGAGAGAGAG	660
OY	661	CCACTTTTATTTTCCAAATACATTCAGTTTACCCAGAAAGCAACGAAAGAAATTAACAA	720
Dp	661	CCACTTTTATTTTCCAAATACATTCAGTTTACCCAGAAAGCAACGAAAGAAATTAACAA	720
OY	721	GGCCAGATTTACCTTATGAGCAGAGAGAGATCAGCTGGACGAGTACAGCCATCCCA	780
Dp	721	GGCCAGATTTACCTTATGAGCAGAGAGATCAGCTGGACGAGTACAGCCATCCCA	780
OY	781	CTCAGTCAAAAGCTACCCAACCATCATTTTCACACAGTCCGCCCAAAACAGAAAGCAAGCAGCTC	840
Dp	781	CTCAGTCAAAAGCTACCCAACCATCATTTTCACACAGTCCGCCCAAAACAGAAAGCAAGCAGCTC	840
OY	841	CTCAGTTATGATTCCTTATACAGATCTGTGGACCGACAGACGCCGTTCGAAATCCAGGGA	900
Dp	841	CTCAGTTATGATTCCTTATACAGATCTGTGGACCGACAGACGCCGTTCGAAATCCAGGGA	900
OY	901	GTGGGACAGATACAGCTATGGCAGTTCCTACTGAGACTCTGTGGGACAGCTCCAACTCCA	960
Dp	901	GTGGGACAGATACAGCTATGGCAGTTCCTACTGAGACTCTGTGGGACAGCTCCAACTCCA	960
OY	961	ACTGCATCACCTGGGAGGGGCAAAATGGGGAGTTCGAAGATGACAGACCCCTGATGAAGTGG	1020
Dp	961	ACTGCATCACCTGGGAGGGGCAAAATGGGGAGTTCGAAGATGACAGACCCCTGATGAAGTGG	1020
OY	1021	CTCGGCGTGTGGGAGAGAGGAAAAAGCAAACTTAACATATATGATGACAAACTCAGCCGTG	1080
Dp	1021	CTCGGCGTGTGGGAGAGAGGAAAAAGCAAACTTAACATATATGATGACAAACTCAGCCGTG	1080
OY	1081	CACATTCGCTACTACTATGACAAAAATATATGACTAAAGTCAATGATGGTAAAGCGTATGGCT	1140
Dp	1081	CACATTCGCTACTACTATGACAAAAATATATGACTAAAGTCAATGATGGTAAAGCGTATGGCT	1140
OY	1141	ACAAATTTGATTTCCAGGGAATGCGTCAGGCCCTCCAGCTCCACCTCCAGAAATCATCA	1200
Dp	1141	ACAAATTTGATTTCCAGGGAATGCGTCAGGCCCTCCAGCTCCACCTCCAGAAATCATCA	1200
OY	1201	TGTACAAATATCCCAATCAGAGACTCCCTCTACATGAGTTCCTACATGACACACCCCAAGAGA	1260
Dp	1201	TGTACAAATATCCCAATCAGAGACTCCCTCTACATGAGTTCCTACATGACACACCCCAAGAGA	1260
OY	1261	TGAACTTTGTATGCTCCCATATCCCTCGTTTGGCCCGTAACTCATCCAGCTTTTGTGTG	1320
Dp	1261	TGAACTTTGTATGCTCCCATATCCCTCGTTTGGCCCGTAACTCATCCAGCTTTTGTGTG	1320

RESULT 2  
 US-09-902-772-3  
 Sequence 3, Application US/09902772  
 Patent No. US20020164739A1  
 GENERAL INFORMATION:  
 APPLICANT: Chugai Seiyaku kabushiki Kaisha, 5001, Iwamoto et  
 TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
 TITLE OF INVENTION: the Proteins  
 FILE REFERENCE: Chugai seiyaku kabushiki Kaisha 5001  
 CURRENT APPLICATION NUMBER: US/09/902,772  
 CURRENT FILING DATE: 2001-07-12  
 PRIOR APPLICATION NUMBER: US/08/878,177  
 PRIOR FILING DATE: 1997-06-18  
 NUMBER OF SEQ ID NOS: 7  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 3  
 LENGTH: 1528  
 TYPE: DNA  
 ORGANISM: C-eryg gene, chicken DNA  
 US-09-902-772-3

Query Match	93.7%	Score 1356	DB 11	length 1528
Best Local Similarity	94.7%	Pred. No. 0		
Matches 1447	Conservative 0	Mismatches 0	Indels 81	Gaps 1
OY	1	GAATTCGCGAGCAATATTTATTTATGCAATTTATAGCATCAATATCTGTGCACA	60	
Db	1	GAATTCGCGAGCAATATTTATTTATGCAATTTATAGCATCAATATCTGTGCACA	60	
OY	61	TTATGGCAGAGACTATTAAAGAGCATTTACAGTGTGAGTGAAGACAGCTCTGTG	120	
Db	61	TTATGGCAGAGACTATTAAAGAGCATTTACAGTGTGAGTGAAGACAGCTCTGTG	120	
OY	121	AGTGTGCTTACGATCGCCGCCACCTTGGCAAGACAGAAATGACACCTCTCTTCAGTG	180	
Db	121	AGTGTGCTTACGATCGCCGCCACCTTGGCAAGACAGAAATGACACCTCTCTTCAGTG	180	
OY	181	AATATGGGCAAAACATCAAAGATGAGCCCGCGCTTCCACAGACACTGGTTATACAGC	240	
Db	181	AATATGGGCAAAACATCAAAGATGAGCCCGCGCTTCCACAGACACTGGTTATACAGC	240	
OY	241	CCCCGGCCAGAGTTACCATTTAAGATGAGGTGTAACCCCAACACAGTTAATGGGTCAAGCA	300	
Db	241	CCCCGGCCAGAGTTACCATTTAAGATGAGGTGTAACCCCAACACAGTTAATGGGTCAAGCA	300	
OY	301	ATTCACTGATGACTGACGCTGGCCAAAGAGGGAATATGGTTTGCAGTTACACAATG	360	
Db	301	ATTCACTGATGACTGACGCTGGCCAAAGAGGGAATATGGTTTGCAGTTACACAATG	360	
OY	361	TTGGGATGCAATATGSAAGCTACATCTGGAAGAGGAAGCAATTCGCCCTCCAAATATAGCAA	420	
Db	361	TTGGGATGCAATATGSAAGCTACATCTGGAAGAGGAAGCAATTCGCCCTCCAAATATAGCAA	420	
OY	421	CCAATGAACGAAGAGTTATTTGTGCACAGATCCTAGCTTATGGAGCACAGACCATGTAC	480	
Db	421	CCAATGAACGAAGAGTTATTTGTGCACAGATCCTAGCTTATGGAGCACAGACCATGTAC	480	
OY	481	GGCAGTGGCTGGAGTGGGCACTGGAAGGATATGGCTTTCAGACGTGACATCTTGTGTG	540	
Db	481	GGCAGTGGCTGGAGTGGGCACTGGAAGGATATGGCTTTCAGACGTGACATCTTGTGTG	540	





```

OY 554 TGGGAAGAGTTGTGTAATAATGACCAAGATGATCTCCAGAGACTCAGCCGAGTATAA 613
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 769 TGGGAAGAGAACTGTGCAAGATGACCAAGAGAGACTTCCAGAGGCTCAGCCAGTACAA 828
OY 614 CGCAGATATCTCTGTGCACACCTACACTACCTCAGAGAGA -----GAGGAGCCAC 664
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 829 CGCCGACATCTCTCTCAGATCTCAGATCTCAGAGAGAGACTCTCTTCCACATTTGAC 888
OY 665 TTTATTTTCCAAATACATAGTTTACCCGAGAGCAGCAAGAAAGATTAACAAGGCC 724
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 889 TTCAGATGATGTTGATTAAGGCTTACAAACTCTCCAGAGTTTGAATGATGCTAAGAACAC 948
OY 725 AGATTTACCTTATGAGCAAGGAGAGATCAGCTGAGAGATCAGTCAACCTTCCCA---C 781
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 949 AGATTTACCATATAGAGCCCCCAGAGAGATCAGCTGAGAGGCTCAGAGCCAGCCAGGCC 1008
OY 782 TCAGTCAAAAGCTACCAACCATCATCTTCAACAGTGCCCAAAACAGAAAGCAGCGTCC 841
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1009 CCAGTCGAAAGCTCTCAACCATCTCTCCACAGTGCCCAAAAGTGAAGACAGCGTCC 1068
OY 842 TCAGTTAGATCTTATCAGATTTCTTGACCGACAGAGCGCTCTTGAATCCAGGAG 901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1069 TCAGTTAGATCTTATCAGATTTCTTGACCGACCAACAGTAGCGCTTGAATCCAGGAG 1128
OY 902 TGGGACAGATACAGCTATGAGCTTCTACTGAGCTTCTGAGAGAGCTTCAACTCCAA 961
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1129 TGGGACAGATCAGCTTGGAGCTTCTCTCTGAGCTTCTGAGAGAGCTTCAACTCCAG 1188
OY 962 CTGATCAGCTTGGAGAGGACCAATGGGAGTTCAAGTAGACAGACCTGATGAAGTGGC 1021
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1189 CTGATCAGCTTGGAGAGGACCAACGAGGAGTTCAAGTAGAGGATCCCGAGAGTGGC 1248
OY 1022 TCGGAGTGGGAGAGAGAAAGCAAACTAATACATATACATATACAACTCAGCGTGC 1081
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1249 CGGAGGCTGGGAGAGAGAGCAACCAACATGATGATGATGATGATGATGATGATGAT 1308
OY 1082 ACTTGGTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1141
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1309 CCTCCGTTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1368
OY 1142 CAATTTGATTTCCAGGAGATCTGTCAGGCTTCCAGCTTCCAGCTTCCAGATATCAT 1201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1369 CAATTTGATTTCCAGGAGATCTGTCAGGCTTCCAGCTTCCAGCTTCCAGATATCAT 1428
OY 1202 GTACAAATACCATCAGACCTCCCTCAGATGATGATGATGATGATGATGATGATGAT 1261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1429 GTACAAATACCATCAGACCTCCCTCAGATGATGATGATGATGATGATGATGATGAT 1488
OY 1262 GAACCTTGTAGCTCCCATCCCTGCTTGGCTTCCCTCAGATGATGATGATGATGATGAT 1321
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1489 GAACCTTGTAGCTCCCATCCCTGCTTGGCTTCCCTCAGATGATGATGATGATGATGAT 1548
OY 1322 CCTCAATCCATCTGGAATTCACCACTGAGAGCATCTTACCCCAATACCAAGCTGAGC 1381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1549 CCTCAATCCATCTGGAATTCACCACTGAGAGCATCTTACCCCAATACCAAGCTGAGC 1608
OY 1382 TGCTCATATGCTTCCATCTTGGACACTTACTACTAA 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1609 CAGCATATGCTTCTCATCTTGGGACTTACTACTAA 1645

```

```

RESULT 4
US-09-864-761-18410/c
: Sequence 18410, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```

```

: FILE REFERENCE: Aecomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 18410
: LENGTH: 567
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AP000163.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
: OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 9.00e-99
: OTHER INFORMATION: EST_HUMAN HIT: R87572.1, EVALUE 0.00e+00
: OTHER INFORMATION: NT HIT: M17254.1, EVALUE 0.00e+00
: US-09-864-761-18410

```

```

Query Match 26.7% Score 387 DB 10 Length 567;
Best Local Similarity 83.7% Pred. No. 8.5e-117;
Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
OY 896 AGGAGTGGGACAGATCAGCTATGCGAGTTCTACTGAGCTTGTGCGACAGTCCAA 955
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 523 AGGAGTGGGACAGATCAGCTATGCGAGTTCTACTGAGCTTGTGCGACAGTCCAA 464
OY 956 CTCACATGATATCAGCTGGGAGGCGCAATGGGAGTTCAATATGACAGACCTCATGA 1015
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 CTCACATGATATCAGCTGGGAGGCGCAATGGGAGTTCAATATGACAGACCTCATGA 404

```

RESULT 5  
 US-09-864-761-20472/c  
 Sequence 20472, Application US/09864761  
 Patent No. US20020048763A1  
 GENERAL INFORMATION:  
 APPLICANT: Penn, Sharon G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wensheng  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 FILE REFERENCE: Aeomica-X-1  
 CURRENT APPLICATION NUMBER: US/09/864,761  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263.6  
 PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670

	Query Match	Best Local Similarity	24.5%;	Score 354.6;	DB 10;	Length 473;
	Matches	399;	Conservative	0;	Mismatches	74;
					Indels	0;
					Gaps	0;
QY	897	GGGAGTGGGCGAGATACAGCTATGCGAGTTCTCTACTGAGCTTCTGTGCGAGAGCTCCAAAC	956			
Db	473	GGCGAGTGGCGCCAGATCCAGCTTTTGGCAGTTCTCTCTGAGGCTCTCTGCGAGAGCTCCAAAC	414			
QY	957	TTCGAACGTGCATACCTCGGAGGGGACAAATGSGGAGTTCAAGATGAGACAGCCCTGATGAA	1016			
Db	413	TTCAGCTGCATACCTCGGAGGGGACCAACAGGGGAGTTCAAGATGAGAGATGCCGATCCGACGAG	354			
QY	1017	GTGGCTCGGCGCTTTGGGGAGAGAGAGAGAAAACAAACCTTAACATGAACATATGACAAACCTCAGC	1078			
Db	353	GTGGCGCCGGCGCTGGGGAGAGCGGAGAGCAAAACCAACATGAACATGAGTAAGCTCAGC	294			
QY	1077	CGTCCACTTGGCTACTCTATGATGACAAAATATTTATGACTAAAGTTTCATGTTAAACGCTAT	1138			
Db	293	CGCCGCCCTCGGTTTACTACTATGACAAAGAACATCATGACCAAGGTCCATGGGAGAACGGCTTAC	234			
QY	1137	GCCATACAAATTTGATTTTCCACGGAATGCGTACAGGCCCTCCAGCCCTCCAGCCCTCCAGGAATCA	1196			
Db	233	GCCATACAAATTTGATTTTCCACGGAATGCGTACAGGCCCTCCAGCCCTCCAGCCCTCCAGGAATCA	174			
QY	1197	TCCATGTACAAATTCAGCATCAGACCTCCCTCTACATGAGTTCTTACCATGACACACCCCGAG	1256			
Db	173	TCTCTGTACAAATTCAGCATCAGACCTCCCTCTACATGAGTTCTTACCATGACACACCCCGAG	114			
QY	1257	AAGATGAACCTTTGAGCTCCCACTCCCTCGTCTTTGGCCCGTAACCTATCCAGCTTTT	1316			
Db	113	AAGATGAACCTTTGAGCTCCCACTCCCTCGTCTTTGGCCCGTAACCTATCCAGCTTTT	54			
QY	1317	GCTCCGCCCTTAATCATCTACTGGAATTCACCAACTGAGAGCATCTAACCCCAATAC	1369			
Db	53	GCTCCGCCCAATCCATACTGGAATTCACCAACTGAGAGCATCTAACCCCAATAC	1			

RESULT 6

US-09-918-995-23356

; Sequence 23356, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED FROM VARIOUS CDNA LIBRARIES

```
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23356
LENGTH: 467
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(467)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23356

Query Match      21.2%; Score 307; DB 12; Length 467;
Best Local Similarity 83.3%; Pred. No. 2e-90;
Matches 349; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 28 AGCAATTATTAGCATCAATATCTGTATCATATTATGGCAGACATTAAGAGCAT 87
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 49 AACCTTATTACATCAATTAATCTGATTCATTATGCGCAGCACTATTAGAGAGCT 108
QY 88 TATCAGTGTGAGTGAAGAACAGCTCTGTTGAGTGTGCTACGATCGCCCACTTG 147
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 109 TATCAGTGTGAGTGAAGAACAGCTCTGTTGAGTGTGCTACGATCGCCCACTTG 168
QY 148 CAAGAAGAAATATACAGCTCTCTCTCCAGTGAATATGGGCAAACTCAAGATGACC 207
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 169 CTAAAGACAGATATACAGCTCTCTCTCCAGTGAATATGGGCAAACTCAAGATGACC 228
QY 208 CGCGCGTCCCGCAGAGCTGTATACAGCCCCCGCAGAGTACCATTAAGATGG 267
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 229 CAGCGTCCCGCAGAGCTGTATACAGCCCCCGCAGAGTACCATTAAGATGG 288
QY 268 AGTGTAAACCAACAGCTTATGCTCAAGGAATTCACCTGATGCTGACGCTGCCAA 327
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 289 AATGTAACTTACGACAGTGAATGCTCAAGGAATTCACCTGATGCTGACGCTGCCAA 348
QY 328 AAGGAGGAAATATGATAGCTTCAACATGTGGGATGAATATGAGAACATGAG 387
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 349 AAGGAGGAAATATGATAGCTTCAACATGTGGGATGAATATGAGAACATGAG 408
QY 388 AAGGAGGAAATATGATAGCTTCAACATGTGGGATGAATATGAGAACATGAG 446
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 409 AGGAGAGACATGCGCCACCCCAACATGACGACGAGAGGATGCTGCGCA 467

RESULT 7
US-09-923-779-54
Sequence 54, Application US/09923779
Patent No. US20020076721A1
GENERAL INFORMATION:
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
FILE REFERENCE: 210121.553
CURRENT APPLICATION NUMBER: US/09/923, 779
CURRENT FILING DATE: 2001-08-06
NUMBER OF SEQ ID NOS: 155
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 54
LENGTH: 549
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 356, 363, 417, 433, 442, 461, 463, 464, 469, 479, 485, 489,
LOCATION: 537, 545
```

```
OTHER INFORMATION: n = A,T,C or G
US-09-923-779-54

Query Match      19.1%; Score 276 6; DB 10; Length 549;
Best Local Similarity 81.2%; Pred. No. 2.4e-80;
Matches 332; Conservative 0; Mismatches 76; Indels 1; Gaps 1;

QY 1011 GATGAAGTGGCTCGCGCTTGGGAGAGAGCAAGCAACCTAATGATGATGACAAA 1070
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2 GACGAGGTGGCGCCCGCTGGGAGAGAGCAAGCAACCTAATGATGATGATGACAAA 61
QY 1071 CTCAGCCGTGCATCTGCTACTACTATGACAAAATATATGACTAAAGTTGATGAA 1130
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 62 CTCAGCCGTGCATCTGCTACTACTATGACAAAATATATGACTAAAGTTGATGAA 121
QY 1131 CGTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCGCTCCAGCCCTCAGCA 1190
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 122 CGTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCGCTCCAGCCCTCAGCA 181
QY 1191 GAATCATCATATGTAACAATACCCATCAGACCTCCCTTACATGATGATGATGATGAT 1250
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 182 GATCATCTGCTGTAACAATACCCATCAGACCTCCCTTACATGATGATGATGATGAT 241
QY 1251 CCCAGAGATGAACTTTGATGCTCCCATCCCTCTGCTTGGCCGTAACCTATCCAGC 1310
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 242 CCACAGAGATGAACTTTGATGCTCCCATCCCTCTGCTTGGCCGTAACCTATCCAGC 301
QY 1311 TTTTTCGTCCTTATTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 1370
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 302 TTTTTCGTCCTTATTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 361
QY 1371 AGGCTGACAGCTGCTCATATGCTTCCATCT-TGGCACCCTACTACTAA 1418
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 362 ANGCTCCCAACAGCCATATGCTTTCATCTGCGGACACTACTACTAA 410

RESULT 8
US-09-864-761-3705/C
Sequence 3705, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864, 761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180, 312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207, 456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632, 366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236, 359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
```

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 3705  
LENGTH: 420  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AP000021.2  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1  
US-09-864-761-3705

Query Match 17.6%; Score 255; DB 10; Length 420;  
Best Local Similarity 81.9%; Pred. No. 2.8e-73;

Matches 294; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1060 ACTATGACAACCTGACCGCTGCACTTGTCTACTATGACAAATATATATGACTAAG 1119  
DB 420 ACTACGATAGCTGACCGCGGCGCTCGTACTACTATGACAAATATATGACCAAG 361  
QY 1120 TTCTAGGTAACGCTATGCTTACAAATTTGTTCCAGCAATGCTGAGCCCTCCAGC 1179  
DB 360 TCATGGGAAGCGCTAGCGCTACAAATTCGACTCCAGGAGTCCAGGCGCTCCAGC 301  
QY 1180 CTCACCGCTCAGATCATTCATGTAACAATACCATGAGACTCCCTACATGAGTTCCT 1239  
DB 300 CCCACCGCGGAGCTATCTCTGTACAAAGTACCCCTGAGACTCCCTACATGAGTTCCT 241  
QY 1240 ACCATGACACCGCCGCAAGATGAACTTTGTAGCTCCCATCCCTGCTTTGCCGTAA 1299  
DB 240 ATACGCGCCACCGCAAGATGAACTTTGTAGGCGCCCACTCCGAGCCCTCCGTGA 181  
QY 1300 CCTCATCAGCTTTTGTGCTGCGCCCTATGCTCATGCTGGAATTCACCACTGAGGCACT 1359  
DB 180 CACTCTCCATCTTTTGTGCTGCGCCCAATGCTGGAATTCACCACTGAGGCACTAT 121  
QY 1360 ACCCGAATACGAGCGCTGCACTGCTCATATGCTTCCCATCTTGAGGACTACTACTAA 1418  
DB 120 ACCCGAATACGAGCGCTGCACTGCTCATATGCTTCCCATCTTGAGGACTACTACTAA 62

## RESULT 9

US-09-864-761-1653/c  
Sequence 1653, Application US/09864761  
Patent No. US20020048763A1

## GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Neomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 1653  
LENGTH: 454  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AP000163.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95  
US-09-864-761-1653

Query Match 17.6%; Score 255; DB 10; Length 454;  
Best Local Similarity 81.9%; Pred. No. 2.9e-73;

Matches 294; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1060 ACTATGACAACCTGACCGCTGCACTTGTCTACTATGACAAATATATATGACTAAG 1119  
DB 454 ACTACGATAGCTGACCGCGGCGCTCGTACTACTATGACAAATATATGACCAAG 395  
QY 1120 TTCTAGGTAACGCTATGCTTACAAATTTGTTCCAGCAATGCTGAGCCCTCCAGC 1179  
DB 394 TCATGGGAAGCGCTAGCGCTACAAATTCGACTCCAGGAGTCCAGGCGCTCCAGC 335  
QY 1180 CTCACCGCTCAGATCATTCATGTAACAATACCATGAGACTCCCTACATGAGTTCCT 1239  
DB 334 CCCACCGCGGAGCTATCTCTGTACAAAGTACCCCTGAGACTCCCTACATGAGTTCCT 275



```

CURRENT APPLICATION NUMBER US/09/8664.761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 29364
LENGTH: 348
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP001426.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73

```

RESULT 13  
 US-09-864-761-5035/c  
 : Sequence 5035, Application US/09864761  
 Patent No. US20020048763A1  
 GENERAL INFORMATION:  
 APPLICANT: Penn, Sharon G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wensheng  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 FILE REFERENCE: Aecomica X-1  
 CURRENT APPLICATION NUMBER: US/09/864,761  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263,6  
 PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30

```
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 5035
;; LENGTH: 477
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP000021.2
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.95
US-09-864-761-5035

Query Match          11.3%: Score 163.2; DB 10; Length 477;
Best Local Similarity 86.5%: Pred. No. 6.4e-43;
Matches 180; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 448 CAGATCCCTACCTTATGAGCAGACACCATGTACGCGAGTGGCTGGAGTGGCAGTGAAG 507
D 453 CAGATCCCTACCTTATGAGCAGACACCATGTGCGGCGAGTGGCTGGAGTGGCAGTGAAG 394
QY 508 AGATGCTCTTCCAGACGTGACATCTTTGTTCCAGAACATTGATGGGAAAGAGTTGT 567
D 393 AATATGGCTTCCAGACGTGACATCTTTGTTCCAGAACATTGATGGGAAAGAGTGT 334
QY 568 GTAAATGACCAAGAGACTTCCAGAGACTCAGCCGAGCTATTAACGAGATATCTCC 627
D 333 GCAAGATGACCAAGAGACTTCCAGAGACTCAGCCGAGCTATTAACGAGATATCTTC 274
QY 628 TGTACACCTACCTACTACTCTCAGAGAGAG 655
D 273 TCTCACAATCTCCTACTACTCTCAGAGAGAG 246

RESULT 14
US-09-864-761-12798/c
;; Sequence 12798, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Hanzel, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aeomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR APPLICATION NUMBER: 2001-05-23
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
```

```
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 12798
;; LENGTH: 519
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP001426.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.52
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.45
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.55
US-09-864-761-12798

Query Match          11.3%: Score 163.2; DB 10; Length 519;
Best Local Similarity 86.5%: Pred. No. 6.7e-43;
Matches 180; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 448 CAGATCCCTACCTTATGAGCAGACACCATGTACGCGAGTGGCTGGAGTGGCAGTGAAG 507
D 359 CAGATCCCTACCTTATGAGCAGACACCATGTGCGGCGAGTGGCTGGAGTGGCAGTGAAG 300
QY 508 AGATGCTCTTCCAGACGTGACATCTTTGTTCCAGAACATTGATGGGAAAGAGTTGT 567
D 299 AATATGGCTTCCAGACGTGACATCTTTGTTCCAGAACATTGATGGGAAAGAGTGT 240
QY 568 GTAAATGACCAAGAGACTTCCAGAGACTCAGCCGAGCTATTAACGAGATATCTCC 627
D 239 GCAAGATGACCAAGAGACTTCCAGAGACTCAGCCGAGCTATTAACGAGATATCTTC 180
QY 628 TGTACACCTACCTACTACTCTCAGAGAGAG 655
D 179 TCTCACAATCTCCTACTACTCTCAGAGAGAG 152

RESULT 15
US-09-864-761-21761/c
;; Sequence 21761, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Hanzel, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;; FILE REFERENCE: Aeomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR APPLICATION NUMBER: 2001-05-23
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
```



;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 21761  
;; LENGTH: 225  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AP000021.2  
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.95  
;; OTHER INFORMATION: NT HIT: M21535.1, EVALUATE 1.00e-113  
;; OTHER INFORMATION: EST HUMAN HIT: AW948986.1, EVALUATE 1.00e-108  
;; OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUATE 3.00e-37  
US-09-864-761-21761

Query Match 11.2%; Score 162.2; DB 10; Length 225;  
Best Local Similarity 86.5%; Pred. No. 8.4e-43;  
Matches 179; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
QY 448 CAGATCTACGTAATGAGACACAGACATGACGAGTGGCTGAGTGGGCACTGAAG 507  
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
207 CAGATCTACGTAATGAGATGACACATGTCGCGAGTGGCTGAGTGGGCGGTGAAG 148  
QY 508 AGTATGCTCTCCAGAGCTGACATCTGTTGTTCCAGAACATGATGGGAAGATTGT 567  
Db | ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| || ||  
147 AATATGGCTCTCCAGAGCTGACATCTGTTGTTCCAGAACATGATGGGAAGAACTGT 88  
QY 568 GTAATATGACCAAGATGACTCCAGAGACTACGCGAGCTATATACGAGATATCTCC 627  
Db | ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| || ||  
87 GCAAGATGACCAAGAGAGACTTCAGAGGCTCAACCCCACTACAGCCGACATCTCTTC 28  
QY 628 TGTCACTACTACTACTACTAGAGAGA 654  
Db | ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| || ||  
27 TCTCAGATCTCTCACTACTACTAGAGAGA 1

Search completed: July 28, 2003, 12:23:28  
Job time : 400.967 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 09:24:34 ; Search time 2774.91 Seconds  
(without alignments)  
8445.286 Million cell updates/sec

Title: US-09-902-772-1  
Perfect score: 1447  
Sequence: 1 gaattccgcagcaatrat.....gaaagaaagcccaagaaa 1447

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estbta:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estcnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_estl3:\*  
13: gb\_estl4:\*  
14: gb\_estl5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_frod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	485.4	33.5	880	14 B0233264	B0233264 AGENCOURT
2	482.2	33.3	715	12 B0388025	B0388025 602412867
3	475.4	32.9	642	13 B1713036	B1713036 1601608.Y
4	468.6	32.4	701	9 A0136709	A0136709 A0136709
5	468.6	32.4	1089	14 B0212127	B0212127 AGENCOURT
6	462	31.9	634	13 B0489636	B0489636 pgm2n.pk0

7	452.2	31.3	668	10 B0660034	B0660034
8	445.2	30.8	915	12 B0390291	B0390291
9	444.2	30.7	928	12 B0259459	B0259459
10	442	30.5	1045	13 B0456833	B0456833
11	436.2	30.1	675	10 B0639043	B0639043
12	432.6	29.9	629	10 B0577887	B0577887
13	428.8	29.6	629	10 B0442803	B0442803
14	425.6	29.4	592	13 B0088063	B0088063
15	380	26.3	827	9 A0456498	A0456498
16	367.8	25.4	522	12 B0566705	B0566705
17	367.2	25.4	516	12 B0691566	B0691566
18	352.8	24.4	949	13 B1558012	B1558012
19	335	23.2	553	12 B0256864	B0256864
20	317.2	21.9	740	13 B1558530	B1558530
21	313.4	21.7	740	9 A0456506	A0456506
22	313	21.6	1074	14 B0953977	B0953977
23	308	21.3	666	12 B0390572	B0390572
24	304.8	21.1	645	10 B0656030	B0656030
25	298.2	20.6	651	10 B0442803	B0442803
26	297.8	20.6	1245	14 B0212133	B0212133
27	297.4	20.6	524	9 A1794504	A1794504
28	293	20.2	420	13 B1181694	B1181694
29	284.6	19.7	674	10 B0238163	B0238163
30	261.6	18.1	1015	9 A1554108	A1554108
31	259.8	18.0	400	10 B0245360	B0245360
32	256.2	17.7	619	10 A0421310	A0421310
33	241	16.7	512	12 B0831331	B0831331
34	232.6	16.1	327	10 A0209658	A0209658
35	232.4	15.1	634	10 B0655125	B0655125
36	228.6	15.8	465	14 B07572	B07572
37	218.8	15.1	551	9 A0467723	A0467723
38	215.8	14.9	815	9 A0446143	A0446143
39	214.2	14.8	1068	14 B0062565	B0062565
40	211	14.6	520	10 B0233460	B0233460
41	209	14.4	756	9 A0445352	A0445352
42	208	14.4	872	14 B0232274	B0232274
43	201.4	13.9	788	9 A0456869	A0456869
44	200.2	13.8	675	13 B0440515	B0440515
45	198	13.7	874	14 B0215477	B0215477

## ALIGNMENTS

RESULT 1  
B0233264  
LOCUS  
DEFINITION AGENCOURT 7565836 NIH\_MGC\_92 Homo sapiens CDNA IMAGE:6041526  
5' mRNA sequence.  
ACCESSION B0233264  
VERSION B0233264.1 GI:20414664  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 880)  
AUTHORS  
TITLE NIH-MGC http://mgc.nci.nih.gov/  
JOURNAL  
Natlional Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LMU at:  
http://image.llnl.gov  
Plate: LLM13279 row: h column: 07  
High quality sequence stop: 608.  
Location/Qualifiers  
1. .880

FEATURES  
SOURCE

```

"Organism"="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6041526"
/clone_1b="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

```

	a	c	g	t	-
BASE COUNT	250	231	207	189	3 others
ORIGIN					

Query Match	33.5%;	Score 485.4;	DB 14;	Length 880;
Best Local Similarity	84.4%;	Pred. No. 9e-139;		
Matches 546;	Conservative 0;	Mismatches 101;	Indels 0;	Gaps 0;

OY	8	GCAGACGATATATTTTATTTTATTCACATTTATTCAGATCAATATTCCTGATCAATTTATGGC	67
Db	19	GAGAAAGCATATTTATTTATTAACATTTATTTAAACGATCAATTAACCTTGATTCATTTATGGC	78
OY	68	AAGCACTATTAAGGAAGCATTAATACAGTGGTGAAGTAAGAACGACCTCTTTTGAAGTGC	127
Db	79	CAGCACTATTAAGGAAGCCTTATCAGTTGTGTGATGAGACACAGTCGTTGTGAGTGTGC	138
OY	128	CTAGCGATCGCCCACTTTCGCAAAACAGAAATGACAGCTCTCTCTTCCAGTGAATATGG	187
Db	139	CTACGGAAAGCCACACCTCGGCTTAACACAGAGATGACCGGTCCTCTCCAGGACATATGG	198
OY	188	GCAAACATCAAAAGATGAGCCCGCGGTTTCCCGAGAGACTGGTATACACGCCCCCGCG	247
Db	199	ACAGACTCCAAAGATGAGCCCAACCGGTCCTCAGAGAGATTGGTGTCTCAACCCCAAC	258
OY	248	CAGAGTTACCATTAAGATGAGGTATTAACCCAAACAGGTTAATGGGTCAAGATTTACCC	307
Db	259	CAGGCTCACCATCAAAATGGAATGTATACCTTCGCCAGGTGAATGGTCTCAAGGAACCTCTCC	318
OY	308	TGATGACTGCAGCGTGGCCAAAAGGAGGAAATGGTTAGCAGTTGACGAATATGGTTGGAT	367
Db	319	TGATGAAATGCAAGTGTGGGCCAAAAGGGGGAAGATGGTGGGACCCAGACACCGTTGGAT	378
OY	368	GAACTATGGAAGCTACATGGAAGAGAGAGACATATTCGCGCTCCAAATATGACAAACCAATGA	427
Db	379	GAACTACGCGCAGCTACATGAGAGGAAGACACATGTGCACCCCCAAACATGACCAACGAAGA	438
OY	428	ACGAAGAATATTTGTGCCAGCAGATCCCTACGTTATGAGAGACAAAGACATGTACGGAGATG	487
Db	439	GCGCAGAGATTTTCGTGCCAGCAGATCCTACGCTATGAGATTAACAGACCAATGTCCGGCAGTG	498
OY	488	GCTGAGAGGGGACAGTGAAGAGATATGATGCTTCCACAGCGGACATCTTTGTTCAGAA	547
Db	499	GCTGAGAGGGGCGGTGAAGAAATATATAGGCTTCCACAGCGTCAACACATCTTTTATTCAGAA	558
OY	548	CATTGATGGGAAAAGATGTGTAAATGACCAAAAGATGACTTCCAGAGACTACGCGGAG	607
Db	559	CATGATGGGGAAGGAACGTGCAMAATGACCAAGGACGACTTCCAGAGGCTCACCCCGAG	618
OY	608	CTATTAAGCGAATATCCTCTGTCACACACTACATCACTCAACAGAGA	654
Db	619	CTACACGCGGACACCTCTTCTCACAATCTCCACTTCCAGACTCCAGAGAGA	665

[illegible]

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 715)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

CDNA library Preparation: Life Technologies, Inc.  
CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov  
Plate: LLAM10420 row: f column: 18  
High quality sequence stop: 693.

FEATURES	Location/Qualifiers
source	1. .715

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4521233"
/clone_1b="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMVSPORT6; Site_1: NotI
Site_2: SalI; Cloned unidirectionally; oligo dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies
Note: this is a NIH_MGC library."

```

BASE COUNT	206 a	188 c	172 g	149 t
ORIGIN				

Query Match	33.3%	Score 482.2;	DB 12;	Length 715;
Best Local Similarity	84.1%;	Pred. No. 8e-138;		
Matches 544;	Conservative 0;	Mismatches 103;	Indels 0;	Gaps 0;

OY	8	GCGAAGCAATTAATTATTTATGACATTTATTACGATCAATATATCTGTGACTCATTTATGGC	67
Dd	39	GAGAAGCAATAATTATTTATTTAACTATTTATTAACGATCAATTAACCTTGATTCATTTATGGC	98
OY	68	AAGCACTATTAAAGGAAGCATTTATCTAGTGGTAGTGAAMACCAGTCCTTGTTTAGTGC	127
Dd	99	CACCACTATTAAAGGAAGCCCTTATCAGTTGTGAGTAGAGCACAGTCGTTGTGTGATGTGC	158
OY	128	CTACGATTCGCCCCCACCTTGCAAGAACAATAAGACAGCCCTCTCTTCCAGTGAATATGG	187
Dd	159	CTACGGAAGCCCAACACTGGCTGAACAAGATGACCCGCTCTCTTCACAGCCACTATGG	218
OY	188	GCAAACATCAAAAGATAGAGCCCCGGGGTTCGCCAGAGACACTGGTTATCACAGCCCCGGC	247
Dd	219	ACGACATTTCCAAGATAGAGCCACGGGTCCTCAGAGGATTTGGTGTCTCAACCCCCAGC	278
OY	248	CAGAGTTACCATTAAAGATGAGTGTATACCCAAACCCAGTTATGTGGTCAAGAAATTCACC	307
Dd	279	CAGGCTACCACTTCAAAATGGAATGTATACCTTAGCCAGGTGAATGGCTCAAGAACCTCC	338
OY	308	TGATGACTCAGAGCTGGCAAAAAGAGGGAAAAATGTTAGCATTTACAGATATGTTGGAT	367
Dd	339	TGATGAATCGAGTGTGGCCAAAGGGGGGAATGTTGGGAGCCAGCCACACCGTTGGAT	398
OY	368	GAACATGTGAAGCTACATGGAAGAGACAGCATATTCCTCCAAATATGACAMACATGA	427
Dd	399	GAACCTACGGCAGCTACATGAGAGGAAMACACATGCCACCCCAACATGACAGAAAGA	458
OY	428	ACGAACAGTTATTTGGCCAGCAGATCCTAGTTATGAGAGCACAACCATGTACGGCAGTG	487
Dd	459	GCGCAAGTTATCTGCGCAGCAGATCCTAGCTATGAGAGTCAACACCATGTGGGCGAGTG	518
OY	488	GCCTGAGTGGGAGTGAAGAGATATGGTCTTCCAGACTGGACATCTTGTGTTCCGAA	547
Dd	519	GCTGGAGTGGGAGTGAAGATATAGCCCTTCCAGACTCAACATCTTGTGTTATTCAGAA	578

Oy	548	CATGATGGAGAAAGCTGTGTAAATAACCAAGAAGTACTCTCCAGACAGCGACGCCGAG	607
Db	579	CATCATGGAGGAAGCACTGTGCATTACACCAGAGCAGCTTCACAGAGCGCTTAACCCGAG	638
Oy	608	CTAATACGCGACATATCCTCTCTGTGCACACCTACTACTCTCAGAGAGA	654
Db	639	CTACAAGCGCGACATCCCTCTCTCTCACATCTCCAATTAECTCAGAGAGA	685
RESULT 3	Bt1713036	642 bp	mRNA linear EST 11-MAR-2002
Bt1713036	je01d08.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5085758 5'		
LOCUS	similar to SW:ERG_HUMAN P11308 TRANSCRIPTIONAL REGULATOR ERG ;		
DEFINITION	mRNA sequence.		
ACCESSION	Bt1713036		
VERSION	Bt1713036.1 GI:15688731		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 642)		
	Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Keastner,K.,		
	Lemishka,T., Scearcce,M., Brestelli,J., Gradwohl,G., Clifton,S.,		
	Hillier,L., Marra,M., Page,D., Wylie,T., Martin,J., Blisstein,A.,		
	Schmitt,A., Theising,B., Rittler,E., Ronko,I., Bennett,J., Cardenas		
	,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.,		
	Jackson,Y., and Powers,Y.		
TITLE	Endocrine Pancreas Consortium		
JOURNAL	Unpublished (2000)		
COMMENT	Other.ESTS: je01d08.xl		
	Contact: Douglas Melton, Klaus H. Keastner, & Hiroshi Inoue		
	Endocrine Pancreas Consortium		
	Harvard University, Howard Hughes Medical Institute		
	Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,		
	MA 02138		
	Tel: 617-495-1812		
	Fax: 617-495-8557		
	Email: dmelton@biohp.harvard.edu		
	Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:		
	Washington University Genome Sequencing Center For information on		
	obtaining a clone please contact: Dr. Hiroshi Inoue		
	(hinoue@lm.wustl.edu)		
	Seq primer: -40RP from glbco		
	High quality sequence stop: 473.		
FEATURES	Location/Qualifiers		
SOURCE	1. 642		
	/organism="Homo saplens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5085758"		
	/clone_lib="HR85 islet"		
	/tissue_type="Purified pancreatic islet"		
	/lab_host="DH10B"		
	/note="Organ: Pancreas; Vector: pBluescript SK(-); Site.1:		
	Note: Site_2: XhoI; cDNA made by oligo-dt priming.		
	Size-selected on agarose gel. Average insert size ~1kb. 5'		
	XhoI site was destroyed after directional cloning.		
	Amplified once. Contact information: Hiroshi Inoue, MD,		
	Metabolism Div. (Alan Permutt lab), Washington University		
	School of Medicine, Box 8127, 660 South Euclid Ave., St.		
	Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:		
	314-362-1916, Fax: 314-747-2692."		
BASE COUNT	178 a 171 c 163 g 129 t	1 others	
ORIGIN			
Query Match	32.9%;	Score 475.4;	DB 13; Length 642;
Best Local Similarity	84.0%;	Pred. No. 9.7e-136;	
Matches 536;	Conservative 0;	Mismatches 102;	Indels 0; Gaps 0;
Oy	15	AATAATATTTATTTAGCAATTTTAGCGATCAATAATCTGTACATATGCGAACGACT	74
Db	4	AATAATATTTATTTATTAACATTATTAACGATCAATTAACCTGTATTCATTATGCGACGACT	63

FEATURES	source
Db	64 ATTAAGGAAGCCTTATCATCGTGTGAGTGAACAGAGCCTCTGTTGATGTGACCTACGGA 123
QY	135 TCGCCCCACCTTGGCAAAAGACAGAAATGACACAGCCCTCTCTTCACGTAGATATGAGCAACA 194
Db	124 ACGCCACACCTGGCTTAAGACAGAGATGACCGCGTCCTCTCTCAGCAGCATATGAGCAGACT 183
QY	195 TCAAGATGACCCCGCGCTTCCCGCAGCAGACTGCTTATCAGACAGCCCGCGCAGAGTT 254
Db	184 TCAGAAGAGACCCACAGCGCTCCACAGAGATGGCTGCTCAACCCCAAGCCAGCGGCTC 243
QY	255 ACCATTAGATGAGGTGTAACTCAACACAGGTTATGGGTCAAGAAATTAACCTGATGAC 314
Db	244 ACCATCAAAAAGGAATGTAACTTACCTTACCGAGGTGATGGCTCAAGAACTCTCTGATGAA 303
QY	315 TGCAGCGCTGGCAAAAGAGGAGAAATGTTAGCAGTTTCAGACAAATGTTGGAGTGAACAT 374
Db	304 TGCAGGTGGCCAAAGGCGGGGAGATGTTGGGCGACGCCAAGCACCAGTTGGGATGAACAT 363
QY	375 GGAAGCTTCATGGAAGAGAGACATATTTCCGCTTCCAATATGACAAACCATGAACGAGA 434
Db	364 GGCAGCTTCATGGAAGAGAGAACACATGTCACCCCAAAACATGACACAGAACGAGCCGAGA 423
QY	435 GTTATTTTGGCAGAGATCCTTACGTTTATGAGCAGACAGCCATGATGAGGCGAGTGGAG 494
Db	424 GTTATCGTGCACAGATCTTACGCTTATGAGATCAGACCATGTTGCGGCGAGTGGCTGAG 483
QY	495 TGGCAGATGAAGGAGTATGCTTCCACAGAGCTGGACATCTTGTGTTCCAGAAACATTGAT 554
Db	484 TGGCGGTGGAAGAATATGGCTTCCAGAGCTCAACATCTTGTATTTCCAGAACATCGAT 543
QY	555 GGGAAAGAGTTGTGTAAATATACCAAGATGACTTCCAGAGACTCAGCGGAGCTATATAC 614
Db	544 GGGGAGGAGACGTGTCAAGATGACCCAGAGACACTTCCAGAGAGCTACACCCCGAGTANCA 603
QY	615 GCAGATATCTCTCTGTCCACACCTACACTACACTCTCAGAGA 652
Db	604 GCCGACATCCTTCTCTCAGATCTCAGACTCTCAGAGA 641
RESULT 4	
LOCUS	A0136709 701 bp mRNA linear EST 02-AUG-2002
DEFINITION	A0136709 PLACE1 Homo sapiens cdna clone PLACE1004911 5', mRNA sequence.
ACCESSION	A0136709
VERSION	A0136709.1 GI:10997248
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 701) Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE	HRI human cdna project
JOURNAL	Unpublished (2000)
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomcshri.co.jp HRI human cdna project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cdna library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1..701

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1004911"
/clone_1bp="PLACE1"
/tissue_type="placenta"
/note="Vector:PMELBSFL"

```

Query Match	32.48;	Score 468.6;	DB 9;	Length 701;
Best Local Similarity	84.08;	Pred. No. 1.3e-133;		
Matches 539; Conservative	0;	Mismatches 102;	Indels 1;	Gaps 1;

JOURNAL  
 COMMENT  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: ccgabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 plate: LLAM1307 row: b column: 09  
 High quality sequence stop: 562.  
 Location/Qualifiers  
 1..1089  
 FEATURES  
 source

```
FEATURES      location/qualifiers
source        1. .1089
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6052136"
/clone_id="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

```

Query Match	32.4%	Score 468.6;	DB 14;	Length 1089;
Best Local Similarity	83.68;	Pred. No. 1.6e-133;		
Matches 542; Conservative	0;	Mismatches 105;	Indels 1;	Gaps 1

Db	270	CAGGGTCACCATCAAAATGGAATGTAACCTTAGCCAGGTGTAATGGCTCAGAGAACTCTCC	329
Oy	308	TGATGACTGGAGCGGTGGCAAAAAGAGGGGAAAAATGGTTAGCACTTCAGACAAATGTTGGAT	367
Db	330	TGATGAATGCAAGTGTGGCCAAAAGGGGGGAAGATGGTGGGAGCCACACACCTGTTGGGAT	389
Oy	368	GAACTATGAGAGCTACATGGAAGAGAGACATATTCGCGCTCCAAATAATGACAAACCAATGA	427
Db	390	GAACTACGGCAGCTACATGAGAGAGACACACTGGCACCCCAACATGACACAGAAAGA	449
Oy	428	ACGAGAAGTATTTGTGTCGCCAGAGATCTACGTTATGAGACAGACACATGACGGCAGTG	487
Db	450	GCGCAGAGTTATGCTGCCAGCAGATCTACCGTTATGAGTAGACAGCATGTGGCGGAGTG	509
Oy	488	GCTGGAGTGGCGAGTGAAGAGTAATGCTTCACAGCTGCGACATCTTGTGTTCCAGAA	547
Db	510	GCTGGAGTGGCGGCTGAAAAGTAATATGGCCTTCAGACGTCACAACTCTTGTTATCCGAA	569
Oy	548	CATTGATGGGAAAAGAGTTGTGTAATAATACCAA - GATGACTTCAGAGACTACGCGGA	606
Db	570	CATCGATGGGAANGAACGTGTGCAMATGATACCAAAGAGAGCATCTTCAGANGCTCAACCCA	629
Oy	607	GCATAGCAGATATCTCTGCTGACACACTACACTACTCTA	648
Db	630	GCTACAGCGCGACACTCTCTCTCACTTCACACTTCCACTTACTACTA	671

RESULT	5
LOCUS	BQ212127
DEFINITION	BQ212127 1089 bp mRNA linear EST 02-MAY-2002 AGNCSCURF_7571287 NIH_MGC_92 Homo sapiens cDNA IMAGE:6052136 5' , mRNA sequence.
ACCESSION	BQ212127
VERSION	BQ212127.1 GI:20392058
KEYWORDS	EST.
SOURCE	human. . . .
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 1089)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	

OY		8	GCAACGAATTAATTATTATTAGCAATTTAATCCGATCAATTAFTCTGTGCATCATTTATGGC	67
Db		21	GAAACGCAATAATATTATTATTAACTTTTAAACGATCAATAACTTGATTCATTAATGGC	80
OY		68	AAGCACTATTAAAGAAGCATTTATCAAGTGTGAGTAGAACACCAGTCCTTGTGTTGAGTGC	127
Db		81	CACACTATTAAAGGAAGCCCTTATCGATGTGTGAGTAGAACCCAGTCGTTGTTGAGTGTGC	140
OY		128	CTACGGATGGCCCCCACCTTGTCGAAAGACGAAATGACACCTTCCTTCCAGTGAATTTGG	187
Db		141	CTAGCGAAGGCCCAACCTTGGCTTAAGACAGAGATGACCGCTCTCTCTCCACGACTATGG	200
OY		188	GCAACAATCAAAAGATGAGCCCGCGCTTCCCAGCAGACTGTGTTATCACAGCCCCGGC	247
Db		201	ACGACTTCCAAGATATAGCCCAACGGGTCCCTCANAGATTTGGCTGTCTCAACCCCGAGC	260
OY		248	CAGAGTTACCAITTAAGATGAGTAGTGTATACCCAAACCGAGTTAATGGGTCAAGAAATTAAC	307
Db		261	CAGGGTCACCATCAAAAATGGAATGTATACCTTAGCCAGGATGATGGCTCAAGGAATCTCC	320
OY		308	TGATGACTCAGAGCTGGCGAAAAAGAGAGGAAAAATGTTAGCACTTCACAGATGTGGGAT	367
Db		321	TGATGATGAGATGTGTGGCCCAAAGGGGGGAAGATGTGTGGCAGCCCAACACOCGTTGGAT	380
OY		368	GAACCTTGTGAAGCTTACATGGAAGAGACACATATTCGCCCTCCAAATTTGACAAACCAATGA	427
Db		381	GAACCTTCGGCAGCTACATGTGAGAGAAACACATGCGCACCCCAAACTGACACGAACGA	440
OY		428	ACGAACAGTTATTGTGGCCAGCAGATGCTTACGTTATGTGAGCACACACATGTACGGGAGTG	487
Db		441	GGCGACAGTTATCTGTGCCAGCAGATCTTACCGCTATGGAATACAGACCATGTGGCGCAGTG	500
OY		488	GCTGAGTGGGAGTGAAGAGACTATGTCCTTCCAGACGTGACATCTTGTGTTCCAGAA	547
Db		501	GCTGAGTGGGCGGTGAAGAAATATGGCTTCCAGAGTCMAACTCTGTTATTCCAGAA	560
OY		548	CATTGATGGGAAGAAGTGTGTAAAATGCAACCAAGATGACTTCCAGAGACTCA-CGCGCA	606
Db		561	CATCCATGGGAAGAACTGTGTCAAGATATACCAACAGACACTTCCAGAGGCTCAACCCCCA	620
OY		607	GCATTAACGCAGATATCTCTGTGCACCTTACACTACTCTAGAGAA	654

Db 621 GTCATGCGCCGACATCTCTCTCCATCTCCACTACTACTCAGAGAGA 668

RESULT 6  
LOCUS BM489636  
DEFINITION BM489636 634 bp mRNA linear EST 07-FEB-2002  
pkm2n.pk011.124 Normalized Chicken Breast Muscle, Leg Muscle, and  
Epiphyseal Growth Plate cDNA library (pkm2n) Gallus gallus CDNA  
TRANSCRIPTIONAL REGULATOR ERG p11560754 transcription factor erg  
chicken emb1/CA54404.11 (x77159) ERG [Gallus gallus], mRNA  
sequence.

ACCESSION BM489636  
VERSION BM489636  
KEYWORDS EST  
SOURCE chicken  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 634)  
Cognburn, L.A. and Monsonego-Ornan, E.  
ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and  
Epiphyseal Growth Plate cDNA library, USDA/IRAFs Animal Genome  
Project  
Unpublished (2002)  
JOURNAL Contact: Larry A. Cognburn  
COMMENT University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: cognburnudel.edu, www.chickest.udel.edu.

FEATURES  
Source Location/Qualifiers  
1.634  
/organism="Gallus gallus"  
/strain="Commercial broiler and Ottawa Res. Centre  
strains 90 & 21"  
/db\_xref="taxon:9031"  
/clone="pkm2n.pk011.124"  
/clone\_lib="Normalized Chicken Breast Muscle, Leg Muscle,  
and Epiphyseal Growth Plate cDNA library (pkm2n)"  
/sex="Male and Female"  
/tissue\_type="Breast muscle, leg muscle and epiphyseal  
growth plate"  
/dev\_stage="Breast, leg: Embryo(d19); post-hatch(1d, 1, 3, 5, 7, 9,  
11 weeks); growth plate(1d, 7d, 14d post-hatch)"  
/lab\_host="E. coli EMDH10B"  
/note="Vector: pcwvSPORT6; Library made from equivalent  
pools of total RNA isolated from each tissue (embryonic  
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth  
plate 33.3% of the final RNA pool). Single pass sequencing  
from 5'-end"

BASE COUNT 192 a 164 c 138 g 140 t

ORIGIN

Query Match 31.9%; Score 462; DB 13; Length 634;  
Best Local Similarity 87.2%; Pred. No. 1.4e-13;  
Matches 553; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 418 CAACCAATGAAGAGATTAATGTCAGCAGATCTTACGTTATGAGCAGACCATG 477  
Db 1 CAACCAATGAAGAGATTAATGTCAGCAGATCTTACGTTATGAGCAGACCATG 60  
QY 478 TACGGCATGCTGAGTGGCAGTGAAGAGATGCTTCCAGAGCTGGACATCTTGT 537  
Db 61 TACGGCATGCTGAGTGGCAGTGAAGAGATGCTTCCAGAGCTGGACATCTTGT 120  
QY 538 TGTTCAGACATGATGAGAAAGAGTGTGTAATAATGACCAAGATGACTTCCAGAGAC 597  
Db 121 TGTTCAGACATGATGAGAAAGAGTGTGTAATAATGACCAAGATGACTTCCAGAGAC 180  
QY 598 TCACGCCAGCTATTAACGCAATATCTCTGTCCACACCTACTACTCAGAGAGA--- 654

Db 181 TCAGCCGAGCATATACGACAGATATCTCTCTGTCACACTACTACTCAGAGAGATC 240  
QY 655 ----- 654  
Db 241 CTCCTCCACATTTGACCTTCAGATGATGATTAAGCCCTTACAAAACCTCCACGTTAA 300  
QY 655 ----- GAGAGCCACTTTTATTTTCCAAATCATCAGTTTACCGAG 696  
Db 301 TGCATGCTAGAAACACAGAGAGAGGACCTTTTATTTTCCAAATCATCAGTTTACCGAG 360  
QY 697 AAGCAACCCAAAGATTAACAACAAGCCAGATTTACCTATGAGCAGAGAGATCAG 756  
Db 361 AAGCAACCCAAAGATTAACAACAAGCCAGATTTACCTATGAGCAGAGAGATCAG 420  
QY 757 CGTGCAGCAGTTCACAGCCATCCCTACGTGTAAGAGCTTACCCATCATCTTCAACAG 816  
Db 421 CGTGCAGCAGTTCACAGCCATCCCTACGTGTAAGAGCTTACCCATCATCTTCAACAG 480  
QY 817 TGCCCAAAACAGAAAGACAGCGCTCTCAGTTAGATCCTTATCAGATTTCTGGACGACCA 876  
Db 481 TGCCCAAAACAGAAAGACAGCGCTCTCAGTTAGATCCTTATCAGATTTCTGGACGACCA 540  
QY 877 GCAGCCGCTTTCGAATTCAGAGAGAGTGGCAGATACAGCTATGCGAGTTCTACTGAGC 936  
Db 541 GCAGCCGCTTTCGAATTCAGAGAGAGTGGCAGATACAGCTATGCGAGTTCTACTGAGC 600  
QY 937 TTCTGTGCGACAGCTCCACATCTCCACATGCTATCAG 970  
Db 601 TTCTGTGCGACAGCTCCACATCTCCACATGCTATCAG 634

RESULT 7  
LOCUS BB660034  
DEFINITION BB660034 RIKEN full-length enriched, 13 days embryo lung Mus  
musculus CDNA clone DA30027104 5', mRNA sequence.  
ACCESSION BB660034  
VERSION BB660034.1 GI:16493855  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 668)  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,  
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,  
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
Unpublished (2001)  
TITLE RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
JOURNAL Contact: Yoshihide Hayashizaki  
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsr.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh,  
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E.,  
Matahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura,  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format





						/note="Organ: testis; Vector: PCMW-SPOF6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: This is a NIH-MGC Library."	
BASE COUNT		265 a	242 c	220 g	201 t		
ORIGIN							
Query Match	30.7%	Score 444.2;	DB 12;	Length 928;			
Best Local Similarity	83.5%;	Pred. No. 5.5e-126;					
Matches 540;	Conservative 0;	Mismatches 103;	Indels 4;	Gaps 3;			
OY	8	GCGAACAATTAATTATTAGCAATTAATGACATTAATTCGTATCATCATTATGGC	67				
Db	33	GAGAAGCAATTAATTATTATTACCTTTATTATACATCAATTAACCTTATTCATTATGGC	92				
OY	68	AAGCATTATTAAGGAACCATTTACTGATGATGAAGAACGACAGCTTGTTTAGTGTGC	127				
Db	93	CAGACATTAATTAAGGAACCCATTATCAGTTGGATGAGGACCAGT-CGTGTTTGGATGTGC	151				
OY	128	CTACGGATCGCCCCACCTTGCAAAGACAGAATATACAGCCTCCTTCCTCACATATATGG	187				
Db	152	CTACGGAAAGCGCAACCTGGCTTAAGACAGAGATACCGCGCTCCACGACGACTATGG	211				
OY	188	GCAAAATCAAAAGTGAAGGCCGGCGCTTCCCACACAGACTGTTATCACAGCCCCGGC	247				
Db	212	ACACACTTCCAAGATGAGCCACCGCTCCCTCACAGAGATGCTGTCTCAACCCCGAC	271				
OY	248	CAGAGTTACCATTTAAGATGAGTGTATTAACCAACACAGATTAAATGGTCAAGAAATTCACC	307				
Db	272	CAG-GTCCACCATCAAAATGGAATGTAAACCCATGACAGTGAATGGCTCAAGAAATCTCC	330				
OY	308	TGATGACTGCAGCTGGCCAAAAGAGGAAAAATGTTAGCATTCACAGATTGTTGGAT	367				
Db	331	TGATGAATGCAGTGTGGCCAAAAGGCGGAGATGTGTGGGACGCCACACCGTTGGGAT	390				
OY	368	GAATTAATGGAAGCTACATGGAAGGAAGCAATTCGCGCTCCCAATTAATGCAACCAATGA	427				
Db	391	GAATTAAGGCAAGCTACATGGAAGGAAGCAATTCACACCCCAACATGACACAGAAAGA	450				
OY	428	ACGAAGATTATTTGGCCAGACAGATCCCTAGCTTATGGAAGACACAGACCATTACGGCAGTG	487				
Db	451	GCGAGAGTTATCTGTCGACAGATGCTTACGCTATGGAATGACAGACATGTGCGCGCAGTG	510				
OY	488	GCTGAGTGGGCAATGGAAGAGTATGTGCTTCCAGACGTGACATCTTGTGTTCCAGAA	547				
Db	511	GCTGAGTGGGCGGTGACAGAAATGTGCTTCCAGACGTCAACATCTTGTATTCAGAA	570				
OY	548	CATTGATGGGAAGAGTTGTGTATAATGACCAAAAGATGACTTCCAGAGACTCACGCCGAG	607				
Db	571	CATCGAATGGGAAGAACTGTGCAAGAGACCAAGGACGACTTCCAGAGGCTCACCCC--A	628				
OY	608	CTTAAACGAGATATCTCCCTGTCACACCTACACTTACTCCAGAGAGA	654				
Db	629	GTACAATGCCGACATCTTCTCTCACAATCTCCACTACTCAGAGAGA	675				
RESULT 10							
LOCUS	BM456833	1045 bp	mRNA	linear	EST 05-FEB-2002		
DEFINITION	ABENCOURT_6404041 NIH_MGC_92 Homo sapiens cdna IMAGE:583753						
ACCESSION	BM456833						
VERSION	BM456833.1						
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
TITLE	NIH-MGC http://mgc.nci.nih.gov/						
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)						
	Unpublished (1999)						

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: c9apds-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L1AM12347 row: f column: 10  
High quality sequence stop: 669.

FEATURES

source

Location/Qualifiers

1. 1045  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5583753"  
/clone\_lib="NIH\_MGC\_92"  
/lisse\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: PCMV-SPORE6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

BASE COUNT

300 a 290 c 230 g 225 t

Query Match

Best Local Similarity 76.6%; Score 442; DB 13; Length 1045;  
Pred. No. 2.8e-125; Matches 585; Conservative 0; Mismatches 165; Indels 14; Gaps 3;

81 GAACATTATCATGAGTGAAGACAGCAGTCTGTTAGTGTGCTCCTACGATCGCC 140  
190 GAACCTTATCATGAGTGAAGACAGCAGTCTGTTAGTGTGCTCCTACGATCGCC 249  
141 CACCTGCAAGAGAGATGACAGCCTCTCTCCATGATATGGGCAACATCAAG 200  
250 CACCTGCTAGAGAGATGACAGCCTCTCTCCATGATATGGGCAACATCAAG 309  
201 ATGAGCCCGCGCTTCCCGACAGAGTGTATACAGCCCGCGGCAAGTACCAT 260  
310 ATGAGCCCGCGCTTCCCGACAGAGTGTATACAGCCCGCGGCAAGTACCAT 369  
261 AAGATGAGTGAATCCCAACAGGTTAATGGGTCAAGAAATTCACCTGATGAGC 320  
370 AAATGGAATGAATCCCTACAGGTTAATGGGTCAAGAAATTCACCTGATGAGC 429  
321 GTGGCAAGAGAGAGAAATGTTAGATTCAGATTCAGATTCAGATTCAGATTC 380  
430 GTGGCAAGAGAGAGAAATGTTAGATTCAGATTCAGATTCAGATTCAGATTC 489  
381 TACATGAGAGAGAGATTCATTCGCTCCCAATATGACAAATGAAAGAGATTT 440  
490 TACATGAGAGAGAGATTCATTCGCTCCCAATATGACAAATGAAAGAGATTT 549  
441 GTGGCAAGAGAGATTCATTCGCTCCCAATATGACAAATGAAAGAGATTT 500  
550 GTGGCAAGAGAGATTCATTCGCTCCCAATATGACAAATGAAAGAGATTT 609  
501 GTGGCAAGAGAGATTCATTCGCTCCCAATATGACAAATGAAAGAGATTT 560  
610 GTGGCAAGAGAGATTCATTCGCTCCCAATATGACAAATGAAAGAGATTT 669  
561 GAGTTGTGTAATGACAAAGATGATTCAGAGATTCAGAGATTCAGAGATTC 620  
670 GAGTTGTGTAATGACAAAGATGATTCAGAGATTCAGAGATTCAGAGATTC 729  
621 ATCTCTCTGACACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 680  
730 ATCTCTCTGACACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 789  
681 ACATCAGTTTACCAAGAGAAC-----GCAAGAAATTAACAAGCCAGATTAC 732

Db

790 GATGTGATTAAGGCTTCAACACTCTCCCGGTTAATGATGCTAGAACACCGCATTTAC 849

Qy

733 CTTATGAGCAAGGAGAGATTCAGCGGTGAGAGATTCAGCGGTGAGAGATTC 788

Db

850 CATTTGAGACCCCCCGGAGATTCAGCGGTGAGAGATTCAGCGGTGAGAGATTC 909

Qy

789 AAGGCTACCAACCATCAT--CTTCAACAGTGGCCAAACAGAA 830

Db

910 AAGGCTTCAACACTCTCCCTTCAACAGGCGCCAAACTGAA 953

RESULT 11

BB639043

LOCUS

DEFINITION

musculus cDNA clone A630068021 5', mRNA sequence.

ACCESSION

BB639043

VERSION

BB639043.1 GI:15401323

KEYWORDS

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 675)

Arekawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Komori, H., Kouda, M., Koyama, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M., and Hayashizaki, Y.

RIKEN Mouse ESTs (Arekawa, T., et al. 2001)

Unpublished (2001)

Contact: Yoshinori Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komori, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Komori, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shingawa, A., Alizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arekawa, T., Ishii, Y., and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, 172-186 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

e mouse tissues.

FEATURES

source

Location/Qualifiers

1. 675  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="A630068021"  
/clone\_lib="RIKEN full-length enriched, 3 days neonate





QY 379 GCTACATGAGAGAGACATATTCGCCCTCCMAATATAGAACCAATGAAGAGAGTTA 438  
 Db 320 GCTCATATGAGAGAGACATATTCGCCCTCCMAATATAGAACCAATGAAGAGAGTTA 379  
 QY 439 TTGTGCCAGAGATCTTACCTTATGAGACACAGACCATGTACGCGAGTGGCTGGAGTGG 498  
 Db 380 TCGGCGCAGAGATCTTACCTTATGAGACACAGACCATGTACGCGAGTGGCTGGAGTGG 439  
 QY 499 CAGTGAAGAGATGATGCTTCCAGAGCGGACATCTGTGTTCACCAATGATGAGGA 558  
 Db 440 CGGTGAAGAGATGATGCTTCCAGAGCGGACATCTGTGTTCACCAATGATGAGGA 499  
 QY 559 AAGAGTGTGTAAATGACCAAGATGATCTTCCAGAGCTCAGCGGAGCTATTAACGAG 618  
 Db 500 AGGAAGCTGTGCAAGATGACCAAGTACGACTTCCAGAGGCTCAGCGGAGCTATTAACGAG 559  
 QY 619 ATATCTCTCTGTCACACTTACCTTACCTCAGAGAGA 654  
 Db 560 ACATCTCTCTCTCAGACTTCCAGACTCAGAGAGA 595

RESULT 14  
 LOCUS BM088063 592 bp mRNA linear EST 19-NOV-2001  
 DEFINITION 501436 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION BM088063  
 VERSION BM088063.1 GI:16998691  
 KEYWORDS EST.  
 SOURCE COW.  
 ORGANISM Bos taurus

REFERENCE Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrnenkrug,S.C., Bennett,  
 G.L., Heaton,M.P., Laegreid,M.W., Rohrer,G.A., Chitko-McKown,C.G.,  
 Pettee,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and  
 Keeler,J.W.  
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle  
 JOURNAL Genome Res. 11 (4), 626-630 (2001)  
 MEDLINE 21180013  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt-trimmed with phred  
 v0.980904.e. Vector identified by cross-match with the -minscore 18  
 and -mismatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCAGCAGC  
 Plate: 139 row: D column: 21  
 Seq primer: ATTTAGGTGACACTATAG.  
 FEATURES  
 source Location/Qualifiers  
 1..592  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="MARC 2BOV"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI;  
 Library made from pooled tissue from testis, thymus,  
 semitendinosus muscle, longissimus muscle, pancreas,  
 adrenal, and endometrium."

BASE COUNT 152 a 177 c 159 g 104 t  
 ORIGIN  
 Query Match 29.4%; Score 425.6; DB 13; Length 592;  
 Best Local Similarity 82.4%; Pred. No. 2.6e-120;

Matches 488; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
 QY 52 TTGATCAATATATGCAAGCACTATTAAAGAAATATATGCTAGTGAACACAGT 111  
 Db 1 TTGATTCATTATATGCGACACATATTAAGAAAGCTTATCTGCTGAGAGACAGT 60  
 QY 112 CCTGTGTGAGTGTGCTTACGAGATGCGCCCACTTGAAGACAGAAATGACACCTCT 171  
 Db 61 CACGTGTGAGTGTGCTTACGAGATGCGCCCACTTGAAGACAGAAATGACACCTCT 120  
 QY 172 CTTCAGTAAATATGCGAAACATCAAGATGAGCCCGCGCTTCCAGCAGAGCTGT 231  
 Db 121 CCTCGCGAGCTAATGAGACAGATCAAGATGAGCCCGCGCTTCCAGCAGAGCTGT 180  
 QY 232 TATCAGACCGCCCGCAGAGTTACATTAGATGAGTGTACCAACAGGTTATG 291  
 Db 181 TGTCTCAACCCCGCAGAGGCTCAACATCAAGATGATGTATACCTTACCAAGTGAACG 240  
 QY 292 GGTCAAGAAATTCCTGATGACTGACAGCGTGGCAAAAGAGAAATGTTAGCAGTT 351  
 Db 241 GCTCAAGAAACTCCCTGATGATGAGTGAAGCAATGCGCAAAAGTGGAGAGAGCC 300  
 QY 352 CAGACAAATGTTGGATGATGATGAAAGCTATGAGAAAGAGACCATATTCGCCCTCAA 411  
 Db 301 CGGACACCGTGGGATGATGATGAAAGCTATGAGAAAGAGACCATATTCGCCCTCAA 360  
 QY 412 ATATGACAAACCAATGAACGAAGATTTATGTGCCAGAGATCTTACCTTATGAGCAG 471  
 Db 361 ACATGACCGCAACGACGCGCGCTCATCTCTGAGATCCACCTATGAGTACAG 420  
 QY 472 ACATGACCGCAACGACGCGCGCTCATCTCTGAGATCCACCTATGAGTACAG 420  
 Db 421 ACATGACCGCAACGACGCGCGCTCATCTCTGAGATCCACCTATGAGTACAG 420  
 QY 421 ACATGACCGCAACGACGCGCGCTCATCTCTGAGATCCACCTATGAGTACAG 420  
 Db 421 ACATGACCGCAACGACGCGCGCTCATCTCTGAGATCCACCTATGAGTACAG 420  
 QY 532 TCTTGTGTTCCAGACATTTGATGGAAGAGTTGTGTAAATGACCAAGATGACTTC 591  
 Db 481 TCTTATTTATTCAGAAATATGACGCGAAGAGCTGTCAAGATGACCAAGATGACTTC 540  
 QY 592 AGACATCAGCGCGAGCTATTAAGCATATCTCTGTCACACTTACACTA 643  
 Db 541 AGAGGCTCACCCGAGCTACACGCTGACATCTTCTGTGCGACCTCCACTA 592

RESULT 15  
 LOCUS AJ456498 827 bp mRNA linear EST 22-APR-2002  
 DEFINITION AJ456498 riken1 Gallus gallus cDNA clone 8ml3f2, mRNA sequence.  
 ACCESSION AJ456498  
 VERSION AJ456498.1 GI:20266594  
 KEYWORDS EST.  
 SOURCE chicken.  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 AUTHORS Buerstedde J M.  
 TITLE Gallus gallus bursal lymphocyte EST  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Buerstedde JM  
 Cellular Immunology  
 Heinrich-Pette-Institute  
 Martinistr. 52, 20251 Hamburg, Germany  
 Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.  
 FEATURES  
 source Location/Qualifiers  
 1..827  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"  
 /clone\_lib="8ml3f2"  
 /clone\_lib="riken1"  
 /cell\_type="bursal lymphocyte"  
 /dev\_stage="2-3 weeks Old"  
 /note="CB inbred strain"

BASE COUNT 233 a 236 c 177 g 181 t  
ORIGIN

Query Match 26.3%; Score 380; DB 9; Length 827;

Best Local Similarity 68.1%; Pred. No. 4.2e-106;

Matches 605; Conservative 0; Mismatches 220; Indels 63; Gaps 3;

```
QY 472 ACCATGTACGGCAGTGGCTGGAGTGGAGAGATGCTTTCAGACGCTGACA 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 AACAGTGGCGCATGGGTGGATGGGCCATAAAGAGATGATGATGAGATGACA 62
QY 532 TCTTGTGTTCCAGACATTTGATGGAAAGATTGTGTAATGACCAAGATGCTCC 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 CCACCTCTTCCAAATATGATGATGCAAAAGAGCTCTGCAAAACGACAGATGACTCC 122
QY 592 AGAGCTCAGCGGAGCATATAGCAGATATCCCTGCTACACCTACACTACCTCAG 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 TCCGACACCTCCCTCTACACACAGAGATTTGTTGCTCACCCTCAGTTACTCCAGGG 182
QY 652 AGAGAGAGCCACTTTTATTTTCCAAATACATCATGTTTACCGAAGACAGCAAGAA 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 AAAGTAG---CTACTGCTGTGTACAAATACCTACATCCACAGAGAGCTTCTCAGCTC 239
QY 712 TAACACAGAGCCGATTTACCTTATGACAGAGAGAGATCAGCGTGACGATCACA 771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 TTGCGACCAAGAGAGTCTCTGTTG----- 266
QY 772 GCCATCCCACTCAGTCAAAAGCTACCAACCATCATCTTCAACAGTGCCCAACAGAG 831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 -----CAGGGACACAAAGCTACACAG----- 299
QY 832 ACCAGCGTCTCAGTTAGATCCTTATCATGATTTCTTGACCGACGACGCGCTTGCAC 891
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 AGCAACGCGCTCAGCCAGATCCCTATCAATTCCTGGGGCCACATAGTGTCTTGCCA 359
QY 892 ATCCAGGAGTGGGACATATGAGTATGCGAGTTCTTCTGAGCTTCTGTGCGACACT 951
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 ATCTGGGAGTGGGACATACACTATGACAGTTCCTCTGAGTGTCTGCGACAGTT 419
QY 952 CCACTCCAACTGCATACCTGAGGGGACCAAAATGGGGAGTTCAAGATGACAGACCTG 1011
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 CCAACGCCAGCTGATATACATGGGAGGACCAATGGGAAATTCAGATGACAGACCTG 479
QY 1012 ATGAAGTGGCTGGCGTTGGGGAGAGAGAAAGCAAACTTACATGAACTATGACAAAC 1071
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 ATGAAGTGGGACGCGCTGGGGAGAGAGCAAAAGCAAGCCACATGATATGACAAAGC 539
QY 1072 TCAGCCGTGCACTTCTGCTACTATGACAAAAATATTATGACTAAAGTTGATGAAC 1131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 TGAAGCCAGGCTTCGATCTACTATGATGAAGACATTATGACCAAGTGCATGCAAAA 599
QY 1132 GCTATGCTCAAAATTTGATTTCCAGGAAATCGCTCAGGCCCTCAGGCTCAGCCCTCAG 1191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 GGTATGCTCAAAATTTGATTTCCAGTTCATGCAATGCGCAGGCTCTCAGGCTCATCCACTG 659
QY 1192 AATCATCATGTACAAATATACCATCAGACCTCCCTTACATGAGTTCTTACCATGACACAC 1251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 AATCATCATGTACAAATATACCATCAGATCTCTCTACATGAGCTTCTTACCATGCCACG 719
QY 1252 CCCAGAGATGAACTTTGAGCTCCCATCCCTGCTTGGCCGTAACTCATCCAGCT 1311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 720 AGCAGAGAGTGAATTTGTACCCCAACCCCTTCTTATGCCGTGACATCATCAGTT 779
QY 1312 TTTTGTGCCCCCTAATCATACATGCAATTCACCAACTGGAGGCACT 1359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 780 TCTTGGAGCAGCCTCAGCTTATTTGACCTTCCTGCTGGAGAGCATTT 827
```

Search completed: July 28, 2003, 12:14:58  
Job time : 2776.91 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:14:59 ; Search time 3594.89 Seconds

(without alignments)

3651.115 Million cell updates/sec

Title: US-09-902-772-2

Perfect score: 2444

Sequence: 1 MASTIKALSVSEDSLFE.....IYENTRLPAHNPMSHIGTY 451

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ -p2n.model -DEV-xlh  
-O/cgn2\_1/USPTO.spool/US09902772/runat.23072003\_093657\_14878/app\_query.fasta.1.1294  
-DB-GenEmbl -GFMT-fastap -SUFFIX-rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-Dlosum62 -TRANS-human40.cdi -LIST=45  
-OUTFMT=ptc -NORM-score-pct -THR-MAX=100 -THR-MIN=0 -ALIGN=15 -MODE=LOCAL  
-USER-US09902772-ECGN\_1\_1\_4707-erunat.23072003\_093657\_14878 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb\_da.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2444	100.0	1413	5 AY065661	AY065661 Gallus ga
2	2444	100.0	1447	6 E31253	E31253 Proteins hav
3	2414.5	98.8	1516	5 GGERG	X77159 G.gallus ER
4	2414.5	98.8	1528	6 E31254	E31254 Protein hav
5	2267	92.8	2209	10 AB073080	AB073080 Mus muscu
6	2192	89.7	2012	5 XLAJ4126	AJ224126 Xenopus l
7	2182	89.3	3166	9 HUMERG2	M17254 Human erg2
8	2179	89.2	1808	10 AB031088	AB031088 Rattus no
9	2159.5	88.4	2133	10 AB073079	AB073079 Mus muscu
10	2153.5	88.1	1798	5 XLAJ4125	AJ224125 Xenopus l
11	2151	88.0	2131	10 AB073078	AB073078 Mus muscu
12	1810	74.1	3126	9 HUMERG11	M21535 Human erg p
13	1637	67.0	1436	5 DRE249590	AJ249590 Dario rer
14	1624.5	66.5	1359	9 AY029368	AY029368 Homo sapi
15	1624.5	66.5	2394	9 BC010115	BC010115 Homo sapi
16	1624.5	66.5	2916	6 A36461	A36461 Sequence 2
17	1624.5	66.5	2938	6 AR080101	AR080101 Sequence
18	1624.5	66.5	2938	9 HSHMGFLI	X67001 H.septiens H
19	1624.5	66.5	2957	9 HUMERGBFLI	M89833 Homo sapien
20	1624.5	66.5	2959	9 BC001670	BC001670 Homo sapi
21	1624.5	66.5	6849	12 AY029367	AY029367 Synthetic
22	1621.5	66.3	1673	9 S45205	S45205 Fli-1-Frien
23	1613.5	66.0	1729	10 MMFLI1	X59421 Mouse Fli-1
24	1579.5	64.6	3490	5 CCRNAFLI	Y14773 Coturnix co
25	1576	64.5	4403	5 XLFILG	X66979 X.laevis mr
26	1567.5	64.1	3545	5 CCFITONCO	Y14774 Coturnix co
27	1536.5	62.9	1932	9 HUMERL1A	M83255 Human Fli-1
28	1486.5	60.8	1245	5 AF177538	AF177538 Dario rer
29	1227	50.2	816	10 S66169	S66169 Erg-3-immun
30	1183.5	48.4	762	9 S72621	S72621 EMS...erg 1
31	965	39.5	1431	9 AF327066	AF327066 Homo sapi
32	956.5	39.1	14518	9 AP001423	AP001423 Homo sapi
33	956.5	39.1	100000	9 AP000021	AP000021 Homo sapi
34	956.5	39.1	100000	9 AP000163	AP000163 Homo sapi
35	956.5	39.1	340000	9 AP001731	AP001731 Homo sapi
36	925	37.8	588	9 S72865	S72865 EMS...EMS-e
37	833.5	34.1	66352	9 HSY17293	Y17293 Homo sapien
38	833.5	34.1	187131	9 AP001535	AP001535 Homo sapi
39	833.5	34.1	217382	9 AP001122	AP001122 Homo sapi
40	682	27.9	634	3 SULERG	M61067 Sea urchin
41	664	27.2	2180	3 AY060316	AY060316 Drosophila
42	581.5	23.8	491	3 HDJ311813	AJ311813 Hediste d
43	543.5	22.2	482	3 PCU278692	AJ278692 Perinereid
44	520.5	21.3	169741	9 AC097468	AC097468 Homo sapi
45	519.5	21.3	192968	2 AC127107	AC127107 Rattus no

RESULT 1

#### ALIGNMENTS

AY065661  
LOCUS AY065661 1413 bp mRNA linear VRT 17-DEC-2001  
DEFINITION Gallus gallus erg isoform C-1-1 mRNA, complete cds; alternatively spliced.  
ACCESSION AY065661  
VERSION AY065661.1 GI:17887440  
KEYWORDS  
SOURCE  
ORGANISM Gallus gallus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
REFERENCE  
AUTHORS 1 (bases 1 to 1413)  
Iwamoto, M., Higuchi, Y., Koyama, E., Enomoto-Iwamoto, M., Kurisu, K., Yeh, H., Abrams, W.R., Rosenbloom, J. and Pacifici, M.  
TITLE Transcription factor ERG variants and functional diversification of chondrocytes during limb long bone development  
J. Cell Biol. 150 (1), 27-40 (2000)  
JOURNAL MEDLINE 20351415  
PUBMED 10893354  
REFERENCE  
AUTHORS 2 (bases 1 to 1413)  
Iwamoto, M., Higuchi, Y., Enomoto-Iwamoto, M., Kurisu, K., Koyama, E., Yeh, H., Rosenbloom, J. and Pacifici, M.  
TITLE The role of ERG (ets related gene) in cartilage development  
Osteoarthritis. Cartil. 9 Suppl A, S41-S47 (2001)  
JOURNAL MEDLINE 21535378  
PUBMED 11680687  
REFERENCE  
AUTHORS 3 (bases 1 to 1413)  
Iwamoto, M., Higuchi, Y., Yeh, H. and Pacifici, M.  
TITLE Direct Submission  
Submitted (05-DEC-2001) Oral Anatomy and Developmental Biology, Osaka University Faculty of Dentistry, 1-8 Yamadaoka, Suita, Osaka 565-0871, Japan  
FEATURES  
source location/Qualifiers  
1..1413 /organism="Gallus gallus"  
/db\_xref="taxon:9031"  
37..1392 /note="alternatively spliced; transcription factor"  
/codon\_start=1  
/product="erg isoform C-1-1"  
/protein\_id="AA140889.1"  
/db\_xref="GI:17887441"  
/translation="MASTIKREALSVSEDSOLFECAYGSPHLAKTEKMTSSSEYQCT SKSPVPQDQWLSOPPARVTIMECNPQVNSRNSPDCSAKGVSSSDNYGM NGSYMEKRIIPPMNTTNERVIVPADPTLMSTDHVRQLEAVAVEXGLPDVILF QNIDKELCMTKDDFORLTPSYNADILSHLYREKRGATFIFPTSYPEATORIT TRPDLPYEOARSAMTSHSPHOSKATOPSSVVPTEDDROPOLDYQILGPTSSRLA NPSSGQLOLMQFLELLSDSNSNCITMECTNEFEKTDDEYARRKGEKSPNNAY DKLSRALRYTYDKNINIMTKYNGKRIATKFDHGLAQLOPHRPSSMTKIPSDLPYSS YHAPQKMNENVADHPALPVTSSSFPAAPNPYNSPTGIIYPTIRLPAAHMPSHLGT Y"  
BASE COUNT 424 a 370 c 311 g 308 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,4e-161 Length: 1413  
Score: 2444.00 Matches: 451  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0  
US-09-902-772-2 (1-451) x AY065661 (1-1413)

QY 1 MetAlaSerThrIleuLysGluAlaLeuSerValIValSerGluAspGlnSerLeuPheGlu 20  
Db 37 ATGGCAAGCACTTTAAGAGCAATATACAGTGTGAGTGAAGCAAGCTCTGTTGAG 96  
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
Db 97 TGTGCGTACGAGTACGCCACCTTGCAGAAAGACAGAAATGACACGCTCCTTCACAGTGA 156

QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60  
Db 157 TATGGCAAGCAATCAAGATGAGAGCCCGCGGTTCCTCCAGCAGAGATGTTATCACAGCCC 216  
QY 61 ProAlaArgValThrIleuLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
Db 217 CCGGCAGAGTTCACATTAAAGTAGAGTGTAAACCAACAGGTTAAATGGCTAAGGAAT 276  
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
Db 277 TCACCTGATACCTGCAGCGTGGCAAAAGAGGAAATGGTTGACAGTTCAGACANTGT 336  
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThr 120  
Db 337 GGGATGAACTATGGAAGCTACATGGAAGAAACACATATTCGGCTCCCAATATGACAACC 396  
QY 121 AsnGlnArgArgValIleValProAlaAspProThrLeuPsrPsrThrAspHisValArg 140  
Db 397 AATGAAGAAAGATTAATGTGCCAGCAGATCTTATGAGCAGACACCATGTGACG 456  
QY 141 GlnTyrLeuGlnTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160  
Db 457 CAGTGGCTGAGAGGCGCAGTGAAGAGTATGCTTCCAGACCTGGACATCTGTGTTTC 516  
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
Db 517 CAGAACATTAATGGAAGAGTGTGTAATGACCAAGATGATCTCCAGACATCCAGC 576  
QY 181 ProSerTyrAsnIleAspIleLeuLeuSerHisIleHisTyrLeuArgGluArgValAla 200  
Db 577 CCGAGCTATACCGAGATATCTCTGTCACACCTACCTACCTCAGAGAGAGAGAGGCC 636  
QY 201 ThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrTrpArg 220  
Db 637 ACTTTATTTTCCAAATATACATCAGTTTACCAGAGCAACGCAAGAAATACAAACAG 696  
QY 221 ProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThrSerHisSerHisProThr 240  
Db 697 CCGAGTTTACCTATGAGCAAGCAGAGAGATCAGCTGGCGAGTCCAGCCCTCCACACT 756  
QY 241 GlnSerLysAlaThrGlnProSerSerSerThrValProAsnThrGlnAspGlnArgPro 260  
Db 757 CAGTCAAAAGCTTCCACCATATCTTCCACAGCTGCCAAACAGAACCCAGCGTCTCT 816  
QY 261 GlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280  
Db 817 CAGTTGATGCTTATCAGATTTTGGACCGCAGCAGCGCTTCCAAATCAGAGGAGT 876  
QY 281 GlnGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSerAsn 300  
Db 877 GGGCAGATACAGCTATGCGAGCTTCTTACGAGACTTCTGCGGACAGCTCCAACTCAAC 936  
QY 301 CysIleThrTyrPgluGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAla 320  
Db 937 TGCATACCTCGGAGGACCAAAATGGGAGTTCAATGACAGACCTTATGAGTGGCT 996  
QY 321 ArgArgTyrPgluGlyLysArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAla 340  
Db 997 CGCGCTGGGAGAGAGAAACCAACCTAACATGATGATGACAAACCTCAGCCGTGCA 1056  
QY 341 LeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360  
Db 1057 CTTCGCTACTACTATGACAAATAATATATGACTTAAAGTTCAATGATGTAAGCTGCTAC 1116  
QY 361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerSerMet 380  
Db 1117 AAATTTGATTTCCAGCGAATCGCTCAGCGCTTCCACCTCACCCTCAGAAATCATCCATG 1176  
QY 381 TyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMet 400  
Db 1177 TACAATATACCATCAGACCTCCCTTACATGAGTCTTACCATGACACACCCAGAGATG 1236  
QY 401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAlaAla 420



|||||  
Db 1237 AACCTTGAGCTCCCGCATCCCGCTTGGCCCGCAACCTCATCCAGCTTTTGTCTGCC 1296  
Qy 421 ProAsnProTyrrTPAsnSerProThrglylylyleTyrrProAsnThrArgLeuProAla 440  
Db 1297 CCTAATCATCTACTGGAATTCACCAACTGAGCATCTACCCCAATATACAGAGCTGCACACT 1356  
Qy 441 AlaHisMetProSerHisLeuGlyThrTyrrTyrr 451  
Db 1357 GCTCATATGCTTCCCATCTTGGCACTTACTAC 1389  
RESULT 2  
E31253  
LOCUS E31253  
DEFINITION Protein having cell calcifying inhibitory activity and gene  
encoding the same.  
ACCESSION E31253.1 GI:13025685  
VERSION JP 199075871-A/1.  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 1447)  
AUTHORS Hiroyasu, I., Yoshinobu, H., Marjio, P., Joel, R. and Helena, E.  
TITLE Protein having cell calcifying inhibitory activity and gene  
encoding the same  
JOURNAL Patent: JP 199075871-A 1 23-MAR-1999;  
CHUGAI PHARMACEUT CO LTD, HE TRUSTEES OF THE UNIVERSITY OF  
PENNSYLVANIA  
COMMENT OS Unidentified  
PN JP 199075871-A/1  
PD 23-MAR-1999  
PF 29-MAY-1998 JP 1998166076  
PR 18-JUN-1997 US 08/878177, 20-JUN-1997 US 60/050297 PI  
HIGROYASU IWAMOTO, YOSHINOBU HIGUCHI, MARJIO PASHIFIKI, PI JOEL  
ROZENBROOM,  
PI HELENA E  
PC C12N15/09, A61K48/00, C07K14/465, C07K16/18, C12Q1/68//A61K38/00,  
PC C12P21/02, C12N15/00, A61K37/02  
PC C12P21/08, C12N15/00, A61K37/02  
CC Strandedness: Single;  
CC Topology: linear;  
FH Key Location/Qualifiers  
FT source 1..1447  
location/Qualifiers  
1..1447  
/organism="Unidentified".  
/db\_xref="taxon:32644"  
BASE COUNT 440 a 374 c 317 g 316 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,44e-161 Length: 1447  
Score: 2444.00 Matches: 451  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-09-902-772-2 (1-451) x E31253 (1-1447)  
Qy 1 MetaLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
Db 63 ATGGCAACACTATTAAAGGAAGCATATACAGTGGAGTGAAGACCACTCTGTGTAG 122  
Qy 21 CysAlaTyrrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40  
Db 123 TGTGCTTACGATGCGCCCACTTGCACAAAGACAGAAATGACAGCCCTCTTCCAGTGA 182  
Qy 41 TyrrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
Db 183 TATGGGCAACATGATAAAGATGAGCCCGCGTCCCGCCAGCAAGACTGTATTACAGCCC 242

Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
Db 243 CCGGCCAGAGTTTACATTAAAGATGAGTGAACCCAAACCGGTTAAAGGTCAAGGAT 302  
Qy 81 SerProAspPAspCysSerValAlaLysGlyLysLysMetValSerSerSerAspAsnVal 100  
Db 303 TCACCTGATGACTGACGCGTGGCAAAAGAGGAGGAAATGTTAGACGTTTACAGCAATGTT 362  
Qy 101 GlyMetAsnTyrrGlySerTyrrMetGluLysHisIleProProProAsnMetThrThr 120  
Db 363 GGGATGAACATATGAGAGCTACATGAGAGAGAGATATTCGCCCTCCAAATATGACAAAC 422  
Qy 121 AsnGluArgValIleValProAlaAspProThrIleuTrpSerThrAspHisValArg 140  
Db 423 AATGACGAAAGATATTGTGCGCAGCATCTTCCGTTATGAGCAGCAGACCATGTACGG 482  
Qy 141 GlnTrpLeuGluTrpAlaValLysGluTyrrGlyLeuProAspValAspIleLeuLeuPhe 160  
Db 483 CAGTGGCTGGAGTGGCGAGTGAAGAGATGTTGTTCCAGACGTGCATCTTGTGTTTC 542  
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
Db 543 CAGAACATTTGATGGGAAAGAGTTGTGTAATAATGACCAAGATGACTTCCAGAGACTCAG 602  
Qy 181 ProSerTyrrAsnAlaAspIleLeuLeuSerHisLeuHisIleTyrrLeuArgGluArgLys 200  
Db 603 CCGAGCTTAATACGAGATATCTCTCTGTCACACTTACCTACGTAGAGAGAGAGACC 662  
Qy 201 ThrPheIlePheProAsnThrSerValTyrrProGlnAlaThrGlnArgIleThrThrArg 220  
Db 663 ACTTTTATTTTCCAAATATCATCACTTTTACCCAGAGACAGCAAGATTAACAACAAG 722  
Qy 221 ProAspLeuProTyrrGluGlnAlaArgSerAlaTrpThrSerHisSerHisProThr 240  
Db 723 CCAGTTTACCTTATGACCAAGCGAGATCAGCGTGAAGAGTCAACAGCAGCATCCCACT 782  
Qy 241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgPro 260  
Db 783 CAGTCAAAAAGCTACCAACCATCATCTTCAACAGGCCCAAAAGAGACAGCGTCTCT 842  
Qy 261 GlnLeuAspProTyrrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280  
Db 843 CAGTTAGATCTTATCATCAATCTTGACACGACCAACCGCTTGGCAATCCAGGAGCT 902  
Qy 281 GlyIleIleGlnLeuTrpGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsn 300  
Db 903 GGGCAGATACAGCTATGAGCATCTCTACTGAGGCTTGTGCGACAGCTCCAACTCAAC 962  
Qy 301 CysIleThrTrpGluGlyThrAsnGlyLysLeuLysMetThrAspProAspGluValAla 320  
Db 963 TGCATACCTCGTGGAGGCGCAAAATGGGAGTTCAAGATGACAGACCTGATGAATGTGCT 1022  
Qy 321 ArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyrrAspLysLeuSerArgAla 340  
Db 1023 CGCGCTTGGGAGAGAGAGAAAGCAACCTTAATGAACTTGAACAACCTACCGCTGCA 1082  
Qy 341 LeuArgTyrrTyrrAspLysAsnIleMetThrLysValHisGlyLysArgTyrrAlaTyrr 360  
Db 1083 CTTCGCTACTACTATGACAAATAATTTATGACTTAAGTTCAATGGTAAACGGTATGCCCTAC 1142  
Qy 361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerSerMet 380  
Db 1143 AAATTTGATTTCCAGGAAATGCTCAGGCGCTCCAGCTCCAGCTCCGAATATCATCCATG 1202  
Qy 381 TyrrLysTyrrProSerAspLeuProTyrrMetSerSerTyrrHisAlaHisProGlnLysMet 400  
Db 1203 TACAATATCCATGACAGCTCCCTACTAGTACTTCTTCAATGACACACCCCAAGATG 1262  
Qy 401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAlaAla 420  
Db 1263 AACCTTGTAGCTCCCATCCCGCTTGGTGGCCGTAACCTCATCCAGCTTTTGTGCTGCC 1322

QY 421 ProasnpProtyrTPasnsSerProthrglygllyleTyProasnthArgLeupProAla 440  
 |||||||  
 Db 1323 CCTATCATCATCGAATTCACCAACGTGAGCAGCATACCCCAATACAGCGCTCCAGACT 1382  
 |||||||  
 QY 441 AlahisMetProSerHisLeuglyThrTytyr 451  
 |||||||  
 Db 1383 GCTCATATGCTTCCCATCTTGACACCTACTAC 1415  
 |||||||

RESULT 3  
 GSERG  
 LOCUS GSERG 1516 bp mRNA linear VRT 27-APR-1995  
 DEFINITION G.gallus ERG mRNA.  
 ACCESSION X77159  
 VERSION X77159.1 GI:790439  
 KEYWORDS ERG gene.  
 SOURCE chicken.  
 ORGANISM Gallus gallus  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 1516)  
 Dhordain,P., Dewitte,F., Desbiers,X., Stehelin,D. and  
 Duterque-Cogulland,M.  
 Mesodermal expression of the chicken erg gene associated with  
 precartilaginous condensation and cartilage differentiation  
 Mech. Dev. 50 (1), 17-28 (1995)  
 95329425  
 7605748  
 2 (bases 1 to 1516)  
 Duterque-Cogulland,M.  
 Direct submission  
 Submitted (17-JAN-1994) M. Duterque-Cogulland, CNRS UA 1160,  
 Oncologie Molculaire, Institut Pasteur, 1 rue Calmette, 59019  
 Lille, FRANCE

FEATURES  
 source location/Qualifiers  
 1..1516  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"  
 /tissue\_type="spleen"  
 /clone\_id="lambda gt10"  
 /dev\_stage="adult"  
 63..1499  
 /gene="ERG"  
 63..1499  
 /gene="ERG"  
 /codon\_start=1  
 /protein\_id="CAA54404.1"  
 /db\_xref="GI:790440"  
 /translation="MASTIKELSVSESDSLFECAYGSPHLAKTEMTASSSEYGGT  
 SKSPRYQDDWLSPAPARYIKKECPNPNQNSRNSPDDCSVAKGKMWSSSDVGM  
 NYSYMEKEHIPPNNMTNERRVLPADPLMSTDHVROMLENAVKYGLPDVILF  
 ONIDKELCKMTKDDFORLPSTYNAIDILSHLHRLTLPPLHLSDDVDALONSPIL  
 MHARNGTGAFETPNTSVPEATORTITRDILPEOARRSAMTSHSPQSKATOPS  
 STYKPTEDQRQDPIYQILISPTSSRLANPSSGQIQLMQFLLELSDSSNCTMBET  
 NGEKMTDPEVARWGERKSKPMNMDKLSRALRYTDKNIMTKVGRKAYAKFEDH  
 GIAOLQPHPESSMYKPSDLPYMSYNAHPQKMFVAPHPALPVTSSSFFAPNP  
 YNMSPTGIGIPNRLPAHMPSHLGTYY"

BASE COUNT 458 a 392 c 327 g 339 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,73e-159 Length: 1516  
 Score: 2414.50 Matches: 450  
 Percent Similarity: 94.14% Conservative: 0  
 Best Local Similarity: 94.14% Mismatches: 1  
 Query Match: 98.79% Indels: 27  
 DB: 5 Gaps: 1

US-09-902-772-2 (1-451) x GSERG (1-1516)

QY 1 MetLaserThrIlelysgluAlaLeuSerValValSerGluaspGlnSerIleupheGlu 20  
 |||||||

Db 63 ATGCCAACACTATTAAAGAACATATATCAAGTGTGATGAGACAGACCTCTGTTGAG 122  
 QY 21 CysAlaTyrglySerProHisLeuAlaLythrcGluethrAlaSerSerSerglu 40  
 |||||||  
 Db 123 TGTCCTTACGGATCGCCCTTGGCAAGACAGAAATGACAGCTCTCTTCCAGTGAA 182  
 |||||||  
 QY 41 TyrglyGlnThrSerLyMetSerProArgValProGlnGlnAspTyrLeuSerglnPro 60  
 |||||||  
 Db 183 TATGGCAACATCAAGAGATGAGCCCGGGTGTCCACAGACAGACGTGTATCAGACCCC 242  
 |||||||

QY 61 ProAlaArgValThrIleLyMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 |||||||  
 Db 243 CCGGCAGAGTTACATTAAAGATGAGTGAACCAACCAAGTAAATGGGTCAAGGAAT 302  
 |||||||

QY 81 SerProAspAspCysSerValAlaLyGlyLyMetValSerSerSerAspAsnVal 100  
 |||||||  
 Db 303 TCACCTATGACTGCACGTGGCAAAAGAGGAAATGGTTAGCAATTCAGACAAATGTT 362  
 |||||||

QY 101 GlyMetAsnTyrglySerTyMetGluGluLyshHisLeuProProAsnMetThrThr 120  
 |||||||  
 Db 363 GGGATGAACATATGGAACTCATGAGAGAGACATATTCGGCTCCAAATATGACAAACC 422  
 |||||||

QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuprSerThrAspHisValArg 140  
 |||||||  
 Db 423 AATGAACGAAGATATATGTCAGACAGATCTACGTTATGAGACACAGACCATGTACGG 482  
 |||||||

QY 141 GlnTrpLeuGlnTrpAlaValLyGluTyrglyLeuProAspValAspIleLeuLeuPhe 160  
 |||||||  
 Db 483 CAGTGGCTGGAGTGGAGAGTGAAGAGATAGTCTTCCACACGTGGACATCTGTGTTTC 542  
 |||||||

QY 161 GlnAsnIleAspGlyLyGluLeuCysLyMetThrLyAspAspPheGlnArgLeuThr 180  
 |||||||  
 Db 543 CAGAACATGATGGAGAAAGATTGTAAATGACCAAAGATGACTCCAGAGACATCAGC 602  
 |||||||

QY 181 ProSerTyrsnAlaAspIleLeuSerHisLeuHisTyLeuArgGluArg----- 198  
 |||||||

Db 603 CCGAGCTATACGAGATATCTCTGTGTACACCTTACATCACTCAGAGAGACTCTCTTT 662  
 |||||||

QY 198 ----- 198  
 |||||||

Db 663 CCACATTGACTCAGATGATGTTGATAGAGCCCTTACAAACCTCCACGGTTAATGCAT 722  
 |||||||

QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyProGluAla 213  
 |||||||

Db 723 GCTAGAAACACAGAGAGGACGACCTTATTTTCCAAATACATCAGTTTACCAGAGCA 782  
 |||||||

QY 214 ThrGlnArgIleThrThrArgProAspLeuProTyrglnIleAlaArgAsnAlaTrp 233  
 |||||||

Db 783 ACGCAAGATATACACAAAGCCGAGATTACCTTATAGGACAGAGAGATCAGCTGG 842  
 |||||||

QY 234 ThrSerHisSerHisProThrGlnSerLyAlaThrGlnProSerSerSerThrValPro 253  
 |||||||

Db 843 ACGAGTCACAGCCATCCCACTCAGTCAAAAGTACCAACCAATCATCTTCCACAGTCCC 902  
 |||||||

QY 254 LysThrGlnAspGlnArgProGlnLeuAspProTyrglnIleLeuGlyProThrSerSer 273  
 |||||||

Db 903 AAAACAAACAAACAGACCTCTCTCAGTATGATCATTAGATCTTGTGACCCAGCAGAGC 962  
 |||||||

QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeu 293  
 |||||||

Db 963 CGTCTTGCAATATCCAGGAGTGGCAGATACACTATAGGACATGCTTCCACTGGAGCTTCTG 1022  
 |||||||

QY 294 SerAspSerSerAsnSerAsnCysIleThrTyrgluGlyThrAsnGlyGluPheLyMet 313  
 |||||||

Db 1023 TCGGACAGCTCCAACTCACTGCTACCTGGGAGGCGCAAAATGGGAGTTCCAGATG 1082  
 |||||||

QY 314 ThrAspProAspGluValAlaArgArgTyrglyGluArgLySerLyProAsnMetAsn 333  
 |||||||

Db 1083 ACAGACCCCTGATGAAGTGCTCGCGCTGGGAGAGAGCAAAAGCAAACTACATGAAAC 1142  
 |||||||

QY 334 TyrsAspLySerSerArgAlaLeuArgTyTyTyTyAspLyAsnIleMetThrLyVal 353  
 |||||||

Db 1143 TATGACAACTCAGCGCTGCTCTGCTACTATGACAAATAATATATGACTAAAGTT 1202  
 |||||||

QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373  
 Db 1203 CATGTTAAACGCTATACCTTACAAATTGATTTCCACGGAATGCTCAGGCTTCCAGCT 1262  
 QY 374 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393  
 Db 1263 CACCTTCCAGATCATCATGATGTAACAATACCATGACGCTCCCTACATGATGTTCCATAC 1322  
 QY 394 HisAlaHisProGlnLysMetAsnMetValAlaProHisProProAlaLeuProValThr 413  
 Db 1333 CATGACACACCCCAACATGAACTTGTAGCTCCCAATCCCTTGGCCCTTAAC 1382  
 QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGlyGlyIleTyr 433  
 Db 1383 TCATCCAGCTTTTGGCTGCCCCCTATATCCATACCTGAAATTCACCACTGGAGGCACTTAC 1442  
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 Db 1443 CCCAATACCAAGGCTCCAGCTGCTCATATGCTTCCATCTTGGCAGCACTACTAC 1496  
 RESULT 4  
 E31254  
 LOCUS E31254 1528 bp DNA linear PAT 18-JUN-2001  
 DEFINITION Protein having cell calcifying inhibitory activity and gene  
 encoding the same.  
 ACCESSION E31254  
 VERSION E31254.1 GI:13025686  
 KEYWORDS JP 199075871-A/2.  
 SOURCE unidentified.  
 ORGANISM unclassified.  
 unclassified.  
 REFERENCE 1 (bases 1 to 1528)  
 AUTHORS Hiroyasu, I., Yoshinobu, H., Marijo, P., Joel, R. and Helena, E.  
 TITLE Protein having cell calcifying inhibitory activity and gene  
 encoding the same  
 JOURNAL Patent: JP 199075871-A 2 23-MAR-1999;  
 CHUGAI PHARMACEUT CO LTD, HE TRUSTEES OF THE UNIVERSITY OF  
 PENNSYLVANIA  
 COMMENT OS Unidentified  
 PN JP 199075871-A/2  
 PD 23-MAR-1999  
 PE 29-MAY-1998 JP 1998166076  
 PR 18-JUN-1997 US 08/878177, 20-JUN-1997 US 60/050297 PI  
 HIROYASU IWAMOTO, YOSHINOBU HIGUCHI, MARIJO PASHIPIKI, PI JOEL  
 ROZENBROOM,  
 PI HELENA E  
 PC C12N15/09, A61K48/00, C07K14/465, C07K16/18, C12Q1/68//A61K38/00,  
 PC C12P21/02,  
 PC C12P21/08, C12N15/00, A61K37/02  
 CC Strandedness: Single;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers  
 FT source 1..1528  
 FEATURES Location/Qualifiers  
 source 1..1528  
 1..1528  
 /organism="unidentified"  
 /db\_xref="taxon:32644"  
 BASE COUNT 464 a 395 c 330 g 339 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.75e-159 Length: 1528  
 Score: 2414.50 Matches: 450  
 Percent Similarity: 94.14% Conservative: 0  
 Best Local Similarity: 94.14% Mismatches: 1  
 Query Match: 98.79% Indels: 27  
 DB: 6 Gaps: 1  
 US-09-902-772-2 (1-451) x E31254 (1-1528)  
 QY 1 MetalaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20

Db 63 ATGCACACACTATTAAAGAGCAATTATACAGTGGAGAGGAAGACCACTGTTGGAG 122  
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 Db 123 TGTGCTTACGAGTATCCGCCACCCTTCCAAAGAGAAAGACAGCCTCTCTTCCAGTAA 182  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnPro 60  
 Db 183 TATGGCCAAACATCAAGATGAGCCCGCGCTTCCCGACGAGCTGGTTATCACAGCCC 242  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 243 CCGGCCACAGTTACCATTAAGATGAGGTAAACCAACCAAGTTAATGGGTCAAGAAAT 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 Db 303 TCACCTGATGACTGACGCGGTGGCAAAAGAGGAAATGTTAGCAGTTTCAAGCAATGTT 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGlnGluLysHisIleProProProAsnMetThrThr 120  
 Db 363 GGGATGAACATGGAAGCTTACATGGAAGAAGCAATATTCGCTCCAAATATGACAACC 422  
 QY 121 AsnGluArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140  
 Db 423 AATGACGAAGAGTTATTGTGTCACACAGATCTCTAGTTATGAGACACAGACATGTACGG 482  
 QY 141 GlnThrLeuGluThrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 Db 483 CAGTGCCTGAGATGGGCGAGTGAAGAGATGATGCTTCCAGACGCTGACATCTGTTGTTTC 542  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 543 CAGAACATTTGATGGAAAGAGTGTGTAAATGACCAAGATGACTTCCAGAGACTCAGC 602  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgValArg----- 198  
 Db 603 CCGAGCTTAAACGACGATATCTCTCTGTACACACTACACTACTACAGAGACTCTCTT 662  
 QY 198 ----- 198  
 Db 663 CCACATTTGACTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722  
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213  
 Db 723 GCTAGAAAACAGAGGAGGACCACTTTATTTTCCAAATACATCATGTTTACCAACAAACA 782  
 QY 214 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgSerAlaArg 233  
 Db 783 ACGCAAGATATACACAAAGGCCAGATTTACTTATGACGAAGCAGAGATCAGCGTGG 842  
 QY 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 253  
 Db 843 ACGAGTCCACAGCCATCCACATCAGTCAAGCAAAACCTACCAACATCATCTTCAACAGTGGCC 902  
 QY 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273  
 Db 903 AAAACAGAGAACCAAGCGCTCAGTATGATCTTATGATGATCTTGGACCAACAGCAGCAGC 962  
 QY 274 ArgLeuAlaAsnProGlyLysGlyGlnIleGlnLeuTyrPglInPheLeuLeuGluLeuLeu 293  
 Db 963 CGTCTTGCAAAATCCAGGAGAGTGGCAGATACAGCATAGGCACTTCTACTGAGCTTCTG 1022  
 QY 294 SerAspSerSerAsnSerAsnCysIleThrTyrPglGlyThrAsnGlyLysPheLysMet 313  
 Db 1023 TCGGACAGCTTCAACTCCAACTGATCATCTCGGAGAGGCGCAAAATGGGAGATTCACAGATG 1082  
 QY 314 ThrAspProAspGluValAlaArgArgThrPglGluArgLysSerLysProAsnMetAsn 333  
 Db 1083 ACGAGCCCTGATGAGTGGCTCGGCGTGGGAGAGAGGAAACCAACCTTAACATGTAAC 1142  
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353

Db 1143 TATGACAAACTGACCGCTGACTTGGTACTACTATGACAAATAATTATGACAAAGTT 1202

Qy 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373

Db 1203 CATGTAAACGCTATGGCTACAAATTTGATTTCCACGGAATCGCTCAGGCCCTCAGGCT 1262

Qy 374 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393

Db 1263 CACCTCCAGAAATATCATGTACAAATATACCATCAGACCTCCCTCATGATGATTCCTAC 1322

Qy 394 HisAlaHisProGluLysMetAsnPheValAlaProHisProAlaLeuProValThr 413

Db 1323 CATCACACCCCGCAGAAAGATGAACTTGTACTCCCATCCCTGCTTTCGCCGTAAAC 1382

Qy 414 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleTyr 433

Db 1383 TCATCCAGCTTTTGGTGGCCCTAATCCATACGATGGAATTCACCACTGGAGGATCTAC 1442

Qy 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451

Db 1443 CCCAATACCAAGCTGCACAGCTGCTCATATGCTCCATCTTGGCACCTACTAC 1496

RESULT 5

AB073080 2209 bp mRNA linear ROD 17-OCT-2001

LOCUS DEFINITION Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG, transcript variant 3, complete cds.

AB073080

ACCESSION AB073080.1 GI:16197544

VERSION

KEYWORDS

SOURCE Mus musculus

ORGANISM Mus musculus cDNA to mRNA, clone:Erg-3.

REFERENCE

AUTHORS Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.

TITLE Mus musculus Erg mRNA

JOURNAL Published only in Database (2001)

REFERENCE 2 (bases 1 to 2209)

AUTHORS Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gscc.riken.go.jp, URL: http://hgp.gscc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

FEATURES

SOURCE location/Qualifiers

1..2209

/organism="Mus musculus"

/db\_xref="taxon:10090"

/chromosome="16"

/clone="Erg-3"

1..2209

/gene="Erg"

148..1608

/gene="Erg"

/note="homolog of human ets-related gene (ERG), transcript variant 3"

/codon\_start=1

/protein\_id="BAB69950.1"

/db\_xref="GI:16197545"

/translation="MIQTTPDPAHNIKEALVSSEDSLEFECAGTGPLAKTEMTASS

SSDYGSOTSKSPRYPOODMLSOAPARVTIMKECPNSOVNSRNSPDCSVNKGKMG

SPDTVSGSYMEKEKHPVPMNTTNRRTVIVPADPTLMSDHYRQLEMAVKEYGL

DVDTLRFQNDKEKELKMTDKDQRLTPRSNADLILSHLYLRTPLPHLTSDDVDA

LQNSPRLHARNTGGAAIFPNTSVPEATQRTTRDPLPEPRRSATWGHSHLTPQ

SKAQPSPSAVPTEDQRLQDLPQILGPTSSRLANGSGQIQLMQPLLELSSNS

NCITMGTESEFKMTDDEVARWGEKSKPMNYDKLSALRYLKYKNTLTKYGR

YAYFDEHGIADALQHPHPRESSLYKPYSDPMGSGYAHAPQKNMFVPPHPALPVTSS

SFASPPRPYNNSPGGITYPNTRLPSHMPSHLGIYY"

BASE COUNT 606 a 608 c 545 g 450 t

ORIGIN

Alignment Scores:

Pred. No.: 5,18e-149 Length: 2209

Score: 2267.00 Matches: 421

Percent Similarity: 90.59% Conservative: 12

Best Local Similarity: 88.08% Mismatches: 17

Query Match: 92.76% Indels: 28

DB: 10 Gaps: 2

US-09-902-772-2 (1-451) x AB073080 (1-2209)

Qy 2 AlasertHrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21

Db 172 GCACCTCATATTAAAGAGGCCCTTGCTGAGTGTGACGACGACCGTACTATTGAGTGT 231

Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetHrAlaSerSerSerGluTyr 41

Db 232 GCCTACGGAAAGCCACACCTGGCTAAGACACAGATGACCCGATCTTCCAGTGAAT 291

Qy 42 GlyLinhTrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 61

Db 292 GGCACGACATCCAAAGATGAGTCCAGAGTCCCTCAGCAGAGTGGTGTCAAGCCCA 351

Qy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81

Db 352 GCCAGGTCACCATCAAGATGAGTGAACACCCCTGATGAGTGTGATGATGATGATGAT 411

Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101

Db 412 CCTGATAGTGCACGTGTGAAACAAAGGTGGAGATGGTGGGACCCCGGATCTGTGGG 471

Qy 102 MetAsnTyrGlySerTyrMetGluGluGlnHisIleProPropoAsnMetThrThrAsn 121

Db 472 ATGAGTACGCGACACTCATGAGAGAGAGATGATGATGATGATGATGATGATGATGAT 531

Qy 122 GluArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHisValArgGln 141

Db 532 GAGCGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591

Qy 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161

Db 592 TGGCTGAGTGGGCGGCGGAAAGATATGGCTCTCATATGGACGCTTACTATTTTCAG 651

Qy 162 AsnIleAspGlyLysGluLeuCysLysMetHrLysAspAspPheGlnArgLeuThrPro 181

Db 652 AATATCATGAGGAGAGGAGCTGTGCAAGATGACAAAGATGACTTCCAGCGGCTCAGCCG 711

Qy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg----- 198

Db 712 AGCTACAAATGCCGCAATTTCTCTCACATCTCCATCTCAGAGAGACTCCCTTCCA 771

Qy 198 ----- 198

Db 772 CATCTGACTTCCGATGACGTTGATAGGCTTTACAAACCTCTCCAGCGTTAATGCATGCC 831

Qy 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 214

Db 832 AGAAACACAGGGGGTGCAGCTTTATTTCCCAATATCTTCAGATATCCGAAGCTACG 891

Qy 215 GluArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThr 234

Db 892 CAAGAATTTACACTAGCCAGATTTACCTTATAGAGCTCCAGGAATATAGCTGTGACC 951

Qy 235 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 253

Db 952 GGCACAGCCACCTCACCTCCATCCCAAGCTGCTAGCCATCTCCCTTGCAGTGC 1011

Qy 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273

Db 1012 AAAAAGTAAAGCAGCGCTCTCATGATTCCTTACAGAGATCTGGGACGACGACAGTAC 1071

Qy 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeu 293



```

QY 245 ThrGlnProSerSerThrValProLysThrGlnuSpGlnArgProGlnLeuAspPro 264
Db 975 TCACACCA---TCTACACAGTTCCCAAAACAGAACGCCAGGCCACACTGATGCTT 1031
QY 265 TyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGln 284
Db 1032 TATAGATCTTGGGCCCAACAGACGCCACTTCCAATCCAGAGAGGGGCAATTCAG 1091
QY 285 LeuTrpGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrTrp 304
Db 1092 CTCGGCAATCTTACGTGAGAACTGCTTCGATGATGTTCCAACTCCAACTGATCTTGG 1151
QY 305 GluGlyThrAsnGlyGluPheLeuMetThrAspProAspGluValAlaArgArgTrpGly 324
Db 1152 GAAGAAACCAATGGAGATTTAAAGATGACCGATCTCTGTAAGTGGCAAGACGTTGGGG 1211
QY 325 GluArgLysSerLysProAsnMetAsnTrpAspLysLeuSerArgAlaLeuArgTrpGly 344
Db 1212 GAGAGGAAACCAACCAACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1271
QY 345 TyrAspLysAsnIleMetThrLysValHisGlyLysArgTrpAlaTrpLysPheAspPhe 364
Db 1272 TACATATAAAATATATAGACAAAGTCATGCAAGCGCTATGCTTACAAATTTGATTTT 1331
QY 365 HisGlyIleAlaGlnAlaLeuGlnProHisProProGlnLysSerMetTrpLysTrpPro 384
Db 1332 CATGGATGCTCTCAAGCTCTTCAACCTCATCCCTCAGATCACTAATGTAAGTATGCC 1391
QY 385 SerAspLeuProTrpMetSerSerTrpHisAlaHisProGlnLysMetAsnPheValAla 404
Db 1392 TCACAACTTCATATATATGAGCTCATACATGACATGACATGACATGACATGACATGAC 1451
QY 405 ProHisProProAlaLeuProValThrSerSerSerPhePheAlaAlaProAsnProTrp 424
Db 1452 CCCATCGCCAGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1511
QY 425 TrpAsnSerProThrGlyGlyIleArgProAsnThrArgLeuProAlaAlaHisMetPro 444
Db 1512 TGGATTTACCACTGAGATGATTTATCCAAATATATGCTGCGTGCAGTACGATATGCT 1571
QY 445 SerHisLeuGlyThrTrpTrp 451
Db 1572 TCTCATCTGGAACTTACTAT 1592

RESULT 7
LOCUS HUMERG2 3166 bp ss-RNA linear PRI 08-NOV-1994
DEFINITION Human erg2 gene encoding erg2 protein, complete cds.
ACCESSION M17254
VERSION M17254.1 GI:182186
KEYWORDS erg 2 protein.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3166)
AUTHORS Rao,V.N., Papas,T.S. and Reddy,E.S.
TITLE erg, a human ets-related gene on chromosome 21: alternative
splicing, polyadenylation, and translation
JOURNAL Science 237 (4815), 635-639 (1987)
MEDLINE 87263429
PUBMED 3299708

FEATURES
source location/Qualifiers
1..3166
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="21q22.3"
257..1645
/gene="ERG"
257..1645
/gene="ERG"
/gene="ERG"
/note="erg 2 protein"
/codon_start=1

```

```

/protein_id="AA52398.1"
/db_xref="GI:182187"
/db_xref="GDB:G00-119-684"
/translation="MIOPTDPAAHIREALSVSEDSLEFECAYGPHLAKTEMTASS
SSDGGTSKMSPRVQDDWLSQPARRTIMECPQSVNRSRSPDECSYAKGKMG
SPDVTGNNYSYWEKMEKPPNMTNERNRIVPADPLTMDVROWLEVAVEGYLP
DVNLLFQNDIGKELEKMTKDDQRLRTPSYNAIDILSHLHLYETPPLPTEEDQRLDPA
LONSPRLHARNNDLPEPPRSAMTGHSPRQSKAQPSPSTVPKEDQRLDPA
OIGPSTSSRLANGSGGQIOLMOFLELLESHSNSSCTTGEERKMDQDEVARRW
GERSKRNMYDKLSRLRYRYXKNIMTKYHCKRYAKRPDHCIAQLOPHPEPSLY
KTPSDLEPKSYHHPKMFVAPHALPVTSSFFAAPNPMTNSFTGTYNTRILP
TSHPSHLGYTY"
repeat_region 291..298
repeat_region 509..516
repeat_region 3143..3149
/note="8-bp repeat sequence"
/note="putative"
BASE COUNT 915 a 770 c 768 g 713 t
ORIGIN 227 bp upstream of PvuII site; chromosome 21.

Alignment Scores:
Pred. No.: 6.69e-143 Length: 3166
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 88.67% Mismatches: 24
Query Match: 89.28% Indels: 14
DB: 9 Gaps: 4

US-09-902-772-2 (1-451) x HUMERG2 (1-3166)
QY 2 AlaserThrIleLysGluAlaLeuSerValSerGlnuSpGlnSerLeuPheGluCys 21
Db 281 GCACGCTATATCAAGAGACGCTTATGCTGATGATGATGATGATGATGATGATGATGAT 340
QY 22 AlaTrpGlySerProHisLeuAlaLysTrpGluMetThrAlaSerSerSerGlyTrp 41
Db 341 GCTACAGGAGCGCACACCTGCTGTAAGACAGATGACCGCGCTCTCCACGACTAT 400
QY 42 GlyGlnThrSerLysMetSerProArgValProGlnLysPheLeuSerGlnProPro 61
Db 401 GGACAGACTTCACAGAGAGACCCAGCGGTCTCCAGAGAGATGGCTGCTCAACCCCA 460
QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 461 GCCAGGCTCACCATCAAAATGGAATGATACCTTACGACAGTGAATGGCTCAAGAACTCT 520
QY 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db 521 CCTGATGAAATGACAGTGTGGCCAAAGCGGAGAAATGCTGGCAGCCACCGTGGG 580
QY 102 MetAsnTrpGlySerTrpMetGluGluLysHisIleProProProAsnMetThrTrpAsn 121
Db 581 ATGAACTACGGCAGCTACATGAGAGAGAACCAATGCCACCCCAACATGACACCGAAC 640
QY 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db 641 GACGCGAGATTTATCGTCCAGACAGATCTTACGCTATGAGTACAGACATGTGCGGAC 700
QY 142 TyrLeuGluTrpAlaValLysGlyTrpGlyLeuProAspValAspIleLeuLeuPheGln 161
Db 701 TGGCTGAGAGCGCGGAGAAAGATATGGCTTCCAGACGCTCAACATCTTATTCAG 760
QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuTrpPro 181
Db 761 AACATGATGGAGAGAGACCTGCAAGATGACCAAGAGAGACATTCACAGAGGCTCACCCC 820
QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
Db 821 AGCTACAGCGCCACATCTTCTCTACATCTTCACTACCTCAGAGAGACTCTCTTCCA 880
QY 202 PheIlePheProAsnThrSerValTrpProGlnAlaThrGlnArgIleThrThrArgPro 221
Db 221 PheIlePheProAsnThrSerValTrpProGlnAlaThrGlnArgIleThrThrArgPro 221

```

Db 881 CATTTG-----ACTTCA-----GATGATGTTGATTAAGCCTTACAAACCTCTCCA 925  
 QY 222 -----aspLeuProTyrGluGlnAlaArgSerAlaTrp 233  
 Db 926 CGGTAATGATGCTAGAAACACAAATTTACATATGAGCCCCCAGGAGAGACACCTGG 985  
 QY 234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThral 252  
 Db 986 ACCGGTCACAGGCCACCCACCGCCAGTGGAAAGCTGCTCAACCATCTCTCCACAGTG 1045  
 QY 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272  
 Db 1046 CCCAAACCTGAAGACAGGCTGCTCAGTATGATTCCTATCAATCTTGAGCAACAAGT 1105  
 QY 273 SerArgLeuAlaAspProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeu 292  
 Db 1106 ACCCCCTTGCAAAATCCAGGAGTGCGGAGATCCAGCTTGAGCTCTCCCTGGAGCTC 1165  
 QY 293 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLys 312  
 Db 1166 CTGTGGAGACGCTCAACTCCAGCTGCAATCCAGTGGAAAGCACCACAGGGAGTTCAAG 1225  
 QY 313 MetThrAspProAspGlyValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332  
 Db 1226 ATGAGGATCCCGAGAGAGGTGGCCCGGCGTGGGAGAGCGGAAAGCAACCCAAACATG 1285  
 QY 333 AsnTyrAspLysLeuSerArgAlaLeuArgTrpTyrTyrTyrAspLysAsnIleMetThrLys 352  
 Db 1286 AACTTCGATAGCTAGCCGCCGCCCTCCGTTACTACTATGACAAACATCATCTGACCCAG 1345  
 QY 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372  
 Db 1346 GTCCATGGGAGAGCGGTAGCGCTACAGCAAGTTCGACTCCACGGGATGCCAGCCCTCCAG 1405  
 QY 373 ProHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 392  
 Db 1406 CCCCACCCCGGAGATCTCTCTGCAAGTACCCTTCAGCTCCCTGACATGGGCTCC 1465  
 QY 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaIleProHisProAlaLeuProAla 412  
 Db 1466 TATCGCGCCACCCACCAAGATGACTTTGTGGCGCCGCCACCTCCAGCCCTCCCGTG 1525  
 QY 413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyLysIle 432  
 Db 1526 ACATCTCCAGTTTGTTCGTGCCCCCAACCATCTGGAATTCACCACTGGGGGTATA 1585  
 QY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 Db 1586 TACCCCAACCTAGGCTCCCAACCAAGCATATGCTCTCTGCGGCACTTACTAC 1642  
 RESULT 8  
 AB031088 1808 bp mRNA linear ROD 07-AUG-2001  
 LOCUS  
 DEFINITION Rattus norvegicus VESPI4 mRNA for vascular endothelial cell specific protein 14, complete cds.  
 ACCESSION AB031088  
 VERSION AB031088.1 GI:15128488  
 KEYWORDS vascular endothelial cell specific protein 14.  
 SOURCE Rattus norvegicus  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 REFERENCE 1 (bases 1 to 1808)  
 AUTHORS Aoki,T., Toyoda,H., Nishimoto,S., Tawara,J., Ukai,Y. and Komurasaki,T.  
 TITLE Identification of VESPI4, a vascular endothelial cell specific protein  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1808)  
 AUTHORS Aoki,T., Toyoda,H., Nishimoto,S., Tawara,J., Ukai,Y. and Komurasaki,T.  
 TITLE Direct Submission

JOURNAL Submitted (12-Aug-1999) Tomomi Aoki, Taiho Pharmaceutical Co., Ltd., Medicinal Research Laboratories, Molecular Biology Lab., 1-403, Yoshino-cho, Omiya, Saitama 3308530, Japan (E-mail:s17704@ccm.taiho.co.jp, Tel:048-663-1111, Fax:048-652-7254)  
 FEATURES  
 source Location/Qualifiers  
 1..1808  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /cissue\_type="Liver"  
 1..1808  
 /gene="VESPI4"  
 63..1430  
 /note="VESPI4"  
 /note="similar to human erg1"  
 /codon\_start=1  
 /product="vascular endothelial cell specific protein 14"  
 /protein\_id="BAB62744.1"  
 /db\_xref="GI:15128489"  
 /translation="MASTIKELSVSRDOSLFECAVGNPHLAKTEMTASSSDVGT SKMSRVRVQDMLSPPARVITKMCNRSQVNGNSDECVTGKMGVSPDYGM STGSTEKEHMPPEMTNERNVIVPADPTLWSTDVHQMWEAVKEGGLDVLVLF QNIDKELCKMTKDQFORLTPSYNADILSLHLYRPLPHLSDDVDKALQNSPRL MHRNTDLPEYEPERSWTGSHSPQSKAOPSPSTYPTEDQPOLDPOILGPTS SRLANPGSGOILQMOFLLELSDNSNCITMEGTNGEKKMTDPDEVARRRGERSKP MNVDKLSRALRYDYDKNIMTKRHGRAYKFPDFHGIQALOPHPRESSLYKPSDLP YMGSHTHPQKNFPAHPHPALPVYSSFFPATPNYMNWSPGIGIIPNTRLASHMPSH LGTY"  
 BASE COUNT 510 a 522 c 426 g 350 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,56e-143 Length: 1808  
 Score: 2179.00 Matches: 406  
 Percent Similarity: 91.30% Conserved: 14  
 Best Local Similarity: 88.26% Mismatches: 26  
 Query Match: 89.16% Indels: 14  
 DB: 10 Gaps: 4  
 US-09-902-772-2 (1-451) x AB031088 (1-1808)  
 QY 1 MetaLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 63 ATGGCAGACATATTAAGAGGCGCTTGACAGTGGAGCAAGACACTACTATTAG 122  
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerSerLys 40  
 Db 123 TGTGCTTACGGAACGCCACACCTGCTTAAGACAGAAAGACCGCATCTCTCCAGTGAC 182  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60  
 Db 183 TATGCCAGACATCAACATGATGATCCAGAGTCCACCAACAGAGCTGCTCTCAAGCCC 242  
 QY 61 ProAlaArgValThrIleLysMetGlnCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 243 CCAGCAGAGGTCACCATCAAGATGAGTGAACCACTAGCCAGGTGAACGGTCCAGAAC 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 Db 303 TCACCTGACGATGAGTGTGACCAAAAGGTGGGAAGATGGGCGGACCTGATPACTGTG 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGlnGluLysHisIleProProAsnMetThrThr 120  
 Db 363 GGAATGACCTACGGCAGCTACATGAGAGAGAACCAATGCGCCGCCCAATATGACACAG 422  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140  
 Db 423 AATGAACGACAGATGATGCTCCCTGACATCTCTGTGAGAGACAGACCATGTCGG 482  
 QY 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 Db 483 CAGTGGCTGAGAGTGGGAGTAAGAAATACGCGCTCTCTACAGTGGAGCTCTTATTATT 542

QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 543 CAGAAATATGACGGGAAGAGCTGTGCAGATGACAAAGATGACTTCCAGAGGCTACT 602  
 QY 181 ProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisThrLeuArgGlyAla 200  
 Db 603 CCGAGCTACAAATGCGACATCTCTCTCCATCTCCAGTACCTCAGAGAGACTCCCTT 662  
 QY 201 ThrPheIlePheProAsnThrSerValThrProGlnIleAlaThrGlnArgIleThrThrArg 220  
 Db 663 CCACATCTG-----ACATCC-----GACACCTGTGATAGGCTTTACAAACTCT 707  
 QY 221 Pro-----AspleuProTrpGluGlnAlaAlaArgSerAla 232  
 Db 708 CCAAGGTAAATGACATGCTAGAAACACAGATTTACTTACCTTACCTCCAGAGATCAACC 767  
 QY 233 TrpThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThr 251  
 Db 768 TGGACGGCCACAGCCACCCACCCCTCAGTCCAAACCTCCAGCATCTCCCTCCACA 827  
 QY 252 ValProLysThrGluAspGlnArgProGlnLeuAspProTrpGlnIleLeuGlyProThr 271  
 Db 828 GTGCCCAAAACTGAAGACAGCGCTCCCATGATGATCTTACAGATCTCGGAGCCACC 887  
 QY 272 SerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGlu 291  
 Db 888 AGTAGTCGCTTCTTATCCAGGTAGTGCCAGATCCAGCTGTGGCAGTTCTGCTAGAA 947  
 QY 292 LeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyAlaPhe 311  
 Db 948 CTCTGCTGTACAGCTCCAACTCCAACTGATCATTACCTGGAGAGCCACCAAGGAGTTTC 1007  
 QY 312 LysMetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsn 331  
 Db 1008 AAGATACACAAACCGGATGAGTGCTCGCGCTGGGGGAGGAGCAAGCAACCCCAAC 1067  
 QY 332 MetAsnTrpAspLysLeuSerArgAlaLeuArgTrpTrpTrpAspLysAsnIleMetThr 351  
 Db 1068 ATGACTATGACAAACCTCAGCGCTGCTCCCTACTACTACACAAACATCATGACC 1127  
 QY 352 LysValHisGlyLysArgTrpAlaTrpLysPheAspPheHisGlyIleAlaGlnAlaLeu 371  
 Db 1128 AAGGTGACGAGGAGGAGGCTATGCTCAAGTTGACTTCACGAGGATGCGCAGGCCCTG 1187  
 QY 372 GlnProHisProProGluSerSerMetTrpLysTrpProSerAspLeuProTrpMetSer 391  
 Db 1188 CAGCCCATCCCGCCGAGTGTCTCTGACAACTACCCCTCCGACCTGCGCTACATGGGC 1247  
 QY 392 SerTrpHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuPro 411  
 Db 1248 TCTATACACACCCACCCGAGAGATGAACTTTGTGGCTCCACCCCTCGGCGCTCCCA 1307  
 QY 412 ValThrSerSerSerPhePheAlaAlaProAsnProTrpTrpAsnSerProThrGlyAla 431  
 Db 1308 GTCACATCTTCCAGTTCTTTGTGTACCCGAAACCCATCTGGAATTCGCGCATGGGGGC 1367  
 QY 432 IleTrpProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTrp 451  
 Db 1368 ATCTACCGAACAACTAGGCTCCAGCCAGCCATATGCTCTCACCCTGGGACACTTACTAC 1427

RESULT 9  
 AB073079 2133 bp mRNA linear ROD 17-OCT-2001  
 LOCUS  
 DEFINITION Mus musculus Ery mRNA, mouse homolog of Human ets-related gene ERG,  
 transcript variant 2, complete cds.  
 AB073079  
 AB073079.1 GI:16197542  
 VERSION  
 KEYWORDS  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 cDNA to mRNA, clone:ERG-2.  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Ozawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.  
 TITLE Mus musculus Ery mRNA  
 JOURNAL Published Only in Database (2001)  
 REFERENCE 2 (bases 1 to 2133)  
 AUTHORS Ozawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 FEATURES  
 source location/Qualifiers  
 1..2133  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 /clone="Erg-2"  
 1..2133  
 /gene="Erg"  
 149..1540  
 /note="Erg"  
 /note="homolog of human ets-related gene (ERG), transcript variant 2"  
 /protein\_id="BAB69949.1"  
 /db\_xref="GI:16197543"  
 /translation="MIOTVPPPAHAIKEALSVSPDSIFPCA YGT PHLA KTEM TASS  
 SSDYGT SKSPVPODDLSQAPARTYITMECNPSQVNSRSPBCSYNKGKANG  
 SPDTVMSTGIVKEKHVPPTNTERRYIVADPTLSTDHVKMLVAVAYEGL  
 DVDVLLFONIDGKEICKMKDDFORLPYSYNADILSHLYRELP,PHLSDVDVKA  
 LONSPRLMNAHNRNGAFAFEPNTSVPEAORTTTPAOPSPSAVKEDQRPQDP  
 YOLIGTSPRLANPGSGGLOIMQFLLELSDSSNSNCITMEGNGEFKMDPDEVAR  
 WGERKSKPMNNYDKLSRALRYYYDKIMTRVHGKRYARFDFGIALOPHPRESSL  
 YKPSDLPYMGSYHAHPQKNFSPHPALPYTSSSPFASPNYNSPTGIGTPNTRL  
 PASHMSHLCTTY"  
 BASE COUNT 587 a 579 c 531 g 436 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,55e-141 Length: 2133  
 Score: 2159.50 Matches: 403  
 Percent Similarity: 87.00% Conservative: 12  
 Best Local Similarity: 84.49% Mismatches: 13  
 Query Match: 88.36% Indels: 49  
 DB: 10 Gaps: 2  
 US-09-902-772-2 (1-451) x AB073079 (1-2133)  
 QY 2 AlaSerThrIleLysGluAlaLeuSerValAlaSerGluAspGlnSerLeuPheGlnCys 21  
 Db 173 GCAGCTCANATTAAGAGAGGCGCTTGCAGTTGTGAGCGAGACAGCACTATTGTAGTGT 232  
 QY 22 AlaTrpGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlyTrp 41  
 Db 233 GCGTACGGAAGCGCAACCTGGCTAGACAGAAATACCGCACTCTTCCAGTACTAT 292  
 QY 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61  
 Db 293 GGCACACATCCCAAGATGAGTCCACAGATGCCCTCCACAGAGACTGCTGTCTCAAGCCCA 352  
 QY 62 AlaArgValThrIleLysMetGlnCysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
 Db 353 GCCAGGTACCATCAATGATGAGTGCACACCCCTAGTCAGGTGAATGTTCCAGGAACCTCA 412  
 QY 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101  
 Db 413 CCGATGAGTGGAGTGTGAACAAAGGTGGAGATGGTGAGCGCCGATGACTGGGGG 472  
 QY 102 MetAsnTrpGlySerTrpMetGluGluLysHisIleProProProAsnMetThrTrpAsn 121  
 Db 473 ATGAGCTACGGCACTACTGAGGAGAGCATGTGTCCCTCCCAATATGACCAAAAT 532





Db 486 ATGAGTCACGCTGCTCCACCAAGACTGGCTCTCACAGCTCCATCCAGGGTGACCATC 545  
 Oy LysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAspAspCysSer 86  
 Db 546 AAGATGGAATGACCCCTTACATGATGATGATGATGATGATGATGATGATGATGATGAT 605  
 Oy 87 ValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGlySer 106  
 Db 606 ATAGCAAAAGAGCAAAATGGGTGGAGGTCGATATGTTGGAAATGAATATATGAGCAGC 665  
 Oy 107 TyrMetGluGluLysHisIleProProProAsnMetThrThrAsnGluArgValIle 126  
 Db 666 TACATTAAGAAAAACACATTTCCTCCCAAAACATGACCAACCAATGAGAGAGTAT 725  
 Oy 127 ValProAlaAspProThrIleThrPheSerThrAspHisValArgGlnTyrPheGluTyrPala 146  
 Db 726 GTACCTGCTGACCTTACTTGTGTGAGACAGATATGTCAGACAGTGGCTAGATGGGCA 785  
 Oy 147 ValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLys 166  
 Db 786 ATAAAGAAATACGGCTTCAGATGTCGATGCTGCTCCAAACATGACGAGGAG 845  
 Oy 167 GluLeuLysLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAsp 186  
 Db 846 GAATTAATGACAGATGACCAAGAAAGATTTCATGACTCACGCAACCTACAAATGCTGAT 905  
 Oy 187 IleLeuLeuSerHisLeuHisTyrLeuArgGluArg----- 198  
 Db 906 ATCCCTCTCTCGCATCTTCACTACCTCAGAGAAACTCCTTCACATTGACTTCAGAT 965  
 Oy 199 -----Gly 199  
 Db 966 GATGTTGATAGGCTTACAAAACTCTCCACGGTTAATGATGATGATGATGATGATGATGAT 1025  
 Oy 200 AlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrArgIleThr 219  
 Db 1026 GCCCTCTTATTTCCCAACCTACCGTATTATACAGCAGAAACCAAGATGCCCAAGC 1085  
 Oy 220 ArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrPheSerHisSerHisPro 239  
 Db 1086 AGCGAAGATTTATCATGATGAGCCCTCAAGGATCTGTCGCAAAATGATCTGCAGCA 1145  
 Oy 240 ThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArg 259  
 Db 1146 ---CCATCAAAAGCATCTCAACA---TCTACCAAGCTTCCCAAAACAGAACCCAGG 1199  
 Oy 260 ProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGly 279  
 Db 1200 CCACAGCTAGATCTTATCATGATTTCTGGCCCAACAGACGACCTTGCAAAATCCAGGA 1259  
 Oy 280 SerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnSer 299  
 Db 1260 AGTGGCAGATTCAGCTCTGGCATCTTACTGGAAGCTGCTTGGATAGTTCACACTCC 1319  
 Oy 300 AsnCysIleThrTyrGluGlyThrAsnGlyGluPheLysMetThrAspProAspGluVal 319  
 Db 1320 AACGCAATCTACTGGAAGAGCAACCAATGAGAAATTAAGATGACCGATCCCGAAGAGTG 1379  
 Oy 320 AlaArgArgTyrGluGlyLysSerLysProAsnMetAsnTyrAspLysLeuSerArg 339  
 Db 1380 GCAAGACGTTGGGGGAGAGAGAAAGCAAAACCAACATGAACTATGACAAAGCTAGCGCT 1439  
 Oy 340 AlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAla 359  
 Db 1440 GCACCTGCTACTACTACGATTAATAATATATATGATGATGATGATGATGATGATGAT 1499  
 Oy 360 TyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGluSerSer 379  
 Db 1500 TACAAATTTGATTTTCATGGGATGTCACAGCTCTTCAACCTCATCTCCAGATCAACC 1559  
 Oy 380 MetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLys 399  
 Db 1560 ATGTACAAAGTATCCCTCAGAACTTCATATATGAGTCTCATCATGACATCCACAGAG 1619

Oy 400 MetAsnPheValAlaProHisProAlaLeuProValThrSerSerPhePheAla 419  
 Db 1620 ATGAACTTTGTGGCCCCCATCCGCCAGCTTACTGTCATGCTCAATGCTTTCTTGA 1679  
 Oy 420 AlaProAsnProTyrTyrPheAsnSerProThrGlyGlyIleTyrProAsnThrArgLeuPro 439  
 Db 1680 GCCCTTAATGCATCTACTGGAATTCACCAACTGAGAGTATTTATCCAAATACTCGCTGCCA 1739  
 Oy 440 AlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 Db 1740 GCTAGCCATATGCTCTCATCTTGAACCTACTAT 1775  
 RESULT 11  
 AB073078 2131 bp mRNA linear ROD 17-OCT-2001  
 LOCUS  
 DEFINITION Mus musculus Erg mRNA, mouse homolog of human ets-related gene ERG,  
 transcript variant 1, complete cds.  
 ACCESSION AB073078  
 VERSION AB073078.1 GI:16191716  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus cDNA to mRNA, clone:Erg-1.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS  
 1 Ozawa,R., Noguchi,H., Taylor,T.D., Takeda,T., Hattori,M. and Sakaki,Y.  
 Mus musculus Erg mRNA  
 Published only in Database (2001)  
 2 (bases 1 to 2131)  
 Ozawa,R., Noguchi,H., Taylor,T.D., Takeda,T., Hattori,M. and Sakaki,Y.  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Direct Submission  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
 1-7-22 Suehiro-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail:hattori@rgsc.riken.go.jp, URL:http://hgp.rgsc.riken.go.jp/,  
 Tel:81-45-503-9111, Fax:81-45-503-9170)  
 FEATURES  
 source  
 1. 2131  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 /clone="Erg-1"  
 1. 2131  
 /gene="Erg"  
 148..1536  
 /gene="Erg"  
 /note="homolog of human ets-related gene (ERG), transcript  
 variant 1"  
 /codon\_start=1  
 /protein\_id="BAB69948.1"  
 /db\_xref="GI:16191717"  
 /translation="MTQTVDPDAHKKEALSVSEDOSLFECAYGFPHLAKTEMAS  
 SSDYGGTSSKMPVPOODWLSQAPARVTKMECPQVNGSRSPDECSYNGKGA  
 SPDVMSYSGYEKEKHPVPPNMTNERIVYADPTLWSTDRQMLEAVVEYGL  
 DVDVLEQNDGKELEKMTKDQDFQRLTPSYNADILSHLYLEETPLPHLTSDDVKA  
 LQSPILMARNTDLPYEPBRSAWGHSHLPQSKAAQPSAVPTKTDQRPQDLY  
 QILGPTSSRIANPGSGQIQIMQDLLELSSSSSSKNTTMEGTNGEKMTDPDEVARW  
 GERKSRPNMNTDKLSRALRYDKNTMTVHGKRVAYKDFGIIAQLDPPHESSLY  
 KPSDLPYMGSYHAHPQKNFVSPHPALPVSSSFASPNPWNSTPGTLYPNTPLP  
 ASHMPSHLGYT"  
 BASE COUNT 582 a 590 c 530 g 429 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.05e-141 Length: 2131  
 Score: 2151.00 Matches: 402  
 Percent Similarity: 91.07% Conservative: 16  
 Best local Similarity: 87.58% Mismatches: 27  
 Query Match: 88.01% Indels: 14  
 DB: 10 Gaps: 4





Db 105 CCTCATACGCCGCTGCTCCTTATCCCAAGACAGACATGACTGCTCCGGGACTCAG 164  
 QY 40 GlnTyrLysGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPheSerGln 59  
 Db 165 GATTACGGCAGACGACCAATCATACCCATCCCTCCAGCAAGAGTGAATTAACAG 224  
 QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79  
 Db 225 ---CCGGTGGGGGAAACGTCACCGAGATAT-----GACCATCATCAACGGATCCAGA 275  
 QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99  
 Db 276 GAGTCGCCGGTGGACTCAGCGTGGGAAATGCAATAAATGGTGGGGGAGACCGAGCG 335  
 QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThr 119  
 Db 336 TCTCAGATGAATCTACTGCTACATGATGAGAAAGTCTCCGCTCCCAACATACAG 395  
 QY 120 ThrAsnGluArgValIleValProAlaAspProThrIleuTyrPheSerThrAspHisVal 139  
 Db 396 ACCAATGAGAGAGATCATGCTCCCGCAGACCCGCTCTGTGTGCTCCGGATCAGCTG 455  
 QY 140 ArgGlnTyrLeuGlnTyrPalaValLysGlyLysLeuProAspValAspIleLeu 159  
 Db 456 CGGAGTGTAGTACGTCGGCCATTAAGAGTACGCTCTCAGAGATCGACAGCGGCATG 515  
 QY 160 PheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeu 179  
 Db 516 TTTCACAGCAGACAGTGGAAAAGACTCTGCAGATGAGCAAGACGCTCTCCAGACTC 575  
 QY 180 ThrProSerTyrAsnAlaAspIleLeuLeuSerHisIleuTyrLeuArgGlyLys 199  
 Db 576 ACCGAGCTTATTAACACCGAGCTCTCTCTCATATCTCAATACCTCAGGAAAAGTACG 635  
 QY 200 AlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThr 219  
 Db 636 TCATCGCATATCATAC--AACACGCCATCTCAGCTGACACGCCCTCCGCTGCTGCA 692  
 QY 220 ArgProAspLeuProTyrGluGlnAlaArgArgSerAlaThrPheSerHisSerHisPro 239  
 Db 693 AAGACATGCGCTTATGATGCTGTAGAGGAGGAGGCGGTGTCGAACAACATCGACAGC 752  
 QY 240 ThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArg 259  
 Db 753 GGAAGAAAGCTCTCCAAACAGTGTCTCAGAGCGTGTCCAAAACCCGGATCAGCTCTGT 812  
 QY 260 ProGlnLeuAspProTyrGlnIleLeuGlyProThrSerArgLeuAlaAsnProGly 279  
 Db 813 CCTCAGCGCAATCTCTTATCAGATATATAGACCCACAGCTGCACTTGTCAACCCAGGC 872  
 QY 280 SerGlyGlnIleGlnLeuTyrPheGlnPheLeuGlnLeuLeuSerAspSerSerAsnSer 299  
 Db 873 TCAGGTAGATCCAGCTGTGGAGTCTCTGTGAGCTCTGTCCGACCGCCAAAGCC 932  
 QY 300 AsnCysIleThrTyrPheGluGlyThrAsnGlyLysPheLysMetThrAspProAspGlyVal 319  
 Db 933 GGCTGCATCACTCGGAGAGAACCAACGAGAGTTCAGATGACGAGCCACAGAGAGGTG 992  
 QY 320 AlaArgArgTyrPheGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArg 339  
 Db 993 GCACGGCGCTGGGGCGAAGCAAGACCAACCAACATGAACTGAACTGACGAGCGGCGA 1052  
 QY 340 AlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAla 359  
 Db 1053 GCCCTGGCTACTACTACGACCAAGACATCATGCAAAAGTGCAGCGCAAGCGCTACGCC 1112  
 QY 360 TyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisPropProGluSerSer 379  
 Db 1113 TACAGTTCGACTTCACGCGCATCGCCAGCCCTGCAACCCATCCACGAGTCCACC 1172  
 QY 380 MetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLys 399  
 Db 1173 ATGTAACAATGACCAAGCGAGCTGCGGTAGCTGCTATATACGCCCCACAGCAAG 1232

QY 400 MetAspPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAla 419  
 Db 1233 GTCAACTTCCTCTCTCCATCCCTCGCTGATGCGCTCATCTCTCCAACTTCTGCT 1292  
 QY 420 AlaProAsnProTyrTyrPheSerProThrGlyLysIleTyrProAsnThrArgLeuPro 439  
 Db 1293 CCCCAACCCCATCATCTGAGCTGCCCAACGAGAGTATTATCCCAATCCAGAGCTCCCA 1352  
 QY 440 Ala-----AlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 Db 1353 CGCCATGCCAATTCACAGCTGCTTACACCTTGATGTTACTAC 1397  
 RESULT 14  
 AY029368 1359 bp mRNA linear PRI 09-MAY-2001  
 LOCUS  
 DEFINITION Homo sapiens friend leukemia integration 1 transcription factor  
 (FLI1) mRNA, complete cds.  
 ACCESSION AY029368  
 VERSION AY029368.1 GI:14017402  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1359)  
 AUTHORS Udhil,B.T.S., Kainey,D.R. and Meredith,D.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-APR-2001) Molecular Medicine Unit, Leeds University,  
 Beckett St, Leeds LS9 7TF, United Kingdom  
 FEATURES  
 source  
 1..1359  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 1..1359  
 /gene="FLI1"  
 1..1359  
 /note="FLI1"  
 /note="ETS transcription factor; involved in the  
 pathogenesis of Ewings sarcoma following translocation  
 with EWS gene on chromosome 22; may have oncogenic  
 potential when fused with EWS gene"  
 /codon\_start=1  
 /product="friend leukemia integration 1 transcription  
 factor"  
 /protein\_id="AAK50443.1"  
 /translation="MDGETIKELASVSDQSLPDSATGAANHLPRADMTASGSPDYG  
 /db\_xref="GI:14017403"  
 /translation="MDGETIKELASVSDQSLPDSATGAANHLPRADMTASGSPDYG  
 PKRINLPPOQEWINOPVRVNRVREYDHNMSRESPPDCSVSKCSKLYGGESPNHY  
 NSTADEKNGPPPPNMTTNERRVVPAADPTLMTQDEHROWLEAIKETSLEIDTSTFO  
 NMDGELEKKNKEDFLRATLYNTEVLSLSLSTRESSLLAYWTTSTDDSSRLSYVE  
 DPSYDVRKRWAMNNNSGLNKPPLGAGTITKNTKEORQPPYQILGPTSSRLANP  
 GSGOILMOFLLELLSDSANASCITWEGTGEFKMDPDEVARRMGEKRSKPNMYDK  
 LSRALRYDYKNTMTVYHGRKAYKVFDFHFIQAOLDPHPESSMYRPSISVPSVH  
 AHOOKXNFVPRHSSMPVYSSSFFGAASQYWTSPDTGITYNPNVPRHNPVHVSILGS  
 Y"  
 BASE COUNT 359 a 422 c 333 g 245 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,77e-104 Length: 1359  
 Score: 1624.50 Matches: 306  
 Percent Similarity: 79.43% Conservative: 57  
 Best Local Similarity: 66.96% Mismatches: 83  
 Query Match: 66.47% Indels: 11  
 DB: 9 Gaps: 7  
 US-09-902-772-2 (1-451) x AY029368 (1-1359)  
 QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 1 ATGAGCGGAGCTATTAAAGAGGCTCTCTGCTGTGAGCAGCAGCAGCTCCCTTTGAC 60

QY 21 CysAlaTyrGly----SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39  
 Db 61 TCAGGTCAGGAGGCGAGCCATCTCCCAAGGCGCATGACTGCTCGGGAGTCTT 120  
 QY 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGln 59  
 Db 121 GACTACGGGCGAGCCCAACAGATCAACCCCTCCACACAGCAGAGTGCATCAATCAG 180  
 QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79  
 Db 181 ---CCAGTGGAGGCTCAACGTCACAGGAGAT-----GACCCATGAAATGATCCAGG 231  
 QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99  
 Db 232 GAGTCTCGGTGACTGAGCGGTAGCAAAATGACAGCAAGCTGTGGGCGGAGGAGTCC 291  
 QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---IleProProProAsnMet 118  
 Db 292 AACCCCATGACTACACAGCTATATGACAGAGATGAGCCCTCCCTCCACACATG 351  
 QY 119 ThrThrAsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHis 138  
 Db 352 ACCACACAGAGAGAGAGATGATGTCCTCCGACAGCCACACTGTGACACAGAGCAT 411  
 QY 139 ValArgGlnTyrLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeu 158  
 Db 412 GTGAGCGAATGGTGGAGTGGCGCATAAAGAGTACAGCTGATGAGATGACACATCC 471  
 QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178  
 Db 472 TTTTTCAGAACAGATGAGCGACAGAACTGTGTAATGAAAGAGGAGACTTCCCTCCG 531  
 QY 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLysIleTyrLeuArgGluArg 198  
 Db 532 GCCACACACCTCTTACACAGCAGAGAGTGTGTTCACACCTCAGTACCTACAGGAA--- 588  
 QY 199 GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThr 218  
 Db 589 ---AGTTCACGTGCGCTATATATCAACCTCCACACCGCAATCCTCAGATTGAGT 645  
 QY 219 ThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrThrSerHisSerHis 238  
 Db 646 GTCAAAGAAAGACCTTCTTATGACTCAGTCAAGAGAGAGCTGTGGGCAAAATCATGAT 705  
 QY 239 ProThrGlnSerLysAlaThrGln---ProSerSerSerThrValProLysThrGluAsp 257  
 Db 706 TCTGGCCTCAACAAAGTCTCCCTTGGAGGGGACACAAACGATCAGTAAATACAGAG 765  
 QY 258 GlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 277  
 Db 766 CAAGGCGCCAGCAGATCGGTATCATGATCCTGGGCGCCAGACAGTGGCTGAGCCAAC 825  
 QY 278 ProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGluLeuLeuSerSerSerSer 297  
 Db 826 CTTGGAAGCGGCGAGATCGCTGTGCAATTCCTCTGTGGCTGCTCTCGACAGGCGCC 885  
 QY 298 AsnSerAsnCysIleThrTyrPglGluGlyThrAsnGlyLysPheLysMetThrAspProAsp 317  
 Db 886 AACCCACGCTTATCAGCTGGAGGGAGCCACAGCGGAGTTCAAAATGACGAGCCCGAT 945  
 QY 318 GluValAlaArgArgTyrPglGlyLysSerLysSerLysProAsnMetAsnTyrAspLysLeu 337  
 Db 946 GAGTGGCGCAGGCGCTGGGCGAGCGAAAGCAAGCCCAACATGATTAAGACACAGCTG 1005  
 QY 338 SerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArg 357  
 Db 1006 AGCGGCGCTCTCTTATGATGATTAATAAAATTAATGACCAAAAGTCAAGGCGCAAGA 1065  
 QY 358 TyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlu 377  
 Db 1066 TATGCTTACAATTTGACTTCCACGCGATTGCCAGGCTCTGACGACCATTCGACCGAG 1125

QY 378 SerSerMetTyrLysTyrProSerAspLeuProTyrMetSerTyrHisAlaHisPro 397  
 Db 1126 TCGTCATGTCACAGTACCTTTCGACATCTCTACATGCTCTCTACATGCGCCACAG 1185  
 QY 398 GlnLysMetAsnPheValAlaProHisProProlaLeuProValThrSerSerPhe 417  
 Db 1186 CAGAGGAGGAAGCTTGTCCCTCCATTCATCTTCATGCTGCTGCTCTCCAGCTTC 1245  
 QY 418 PheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyLysIleTyrProAsn----- 435  
 Db 1246 TTTGAGACCGCATCATCATATGAGACTCCCGCCAGGGGGAATCTACCCACAC 1305  
 QY 436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 Db 1306 GTCCCGCGCATCTTAACACCCAGCTGCTTACACATTAGGACACTACTAC 1356  
 RESULT 15  
 BC010115  
 LOCUS BC010115 2394 bp mRNA linear PRI 12-JUL-2001  
 DEFINITION Homo sapiens, clone MGC:19589 IMAGE:3635042, mRNA, complete cds.  
 ACCESSION BC010115  
 VERSION BC010115.1 GI:14603315  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1. (bases 1 to 2394)  
 Strausberg, R.  
 Direct Submission  
 Submitted (02-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/nisc.mgc@hgrl.nih.gov>  
 Contact: Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Binkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCluskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J.,  
 Tlonsgon, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAL Plate: 27 Row: 9 Column: 16  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 257353.  
 Location/Qualifiers  
 1. 2394  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:19589 IMAGE:3635042"  
 /tissue\_type="Lymph. Burkitt lymphoma"  
 /clone\_id="NIH MGC\_8"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOFB7"  
 126..1484  
 /codon\_start=1  
 /product="Unknown (protein for MGC:19589)"  
 /protein\_id="AAH0115.1"  
 /db\_xref="GI:14603316"  
 /translation="MDGTKEALSVSDQSLPDSAYGAANHLPKADMTASGSPGYQ

PHKINLPPOOEMINOPVRYNVRKREYDHMANGSRESVPVDCSVSKSKLGGESNPMMY  
 NSYDEKNGPPPMNTTNERVIVPADPITLMOEVRHMLFMAIKEYSLAMIDISFRO  
 NMDKEKLCNKAKDEFLRATITTYNTEVLSHSTYRESLLANTNTSHDSSRSIYKVE  
 DPSTDSVRKAMGNMNSGLNKSPPLGGAOTISKNEORPPDPQIIGPTSSRIANP  
 GSGQIQIMQFLELLSDSANASCITMETNEDEFTDDEFAWRKGEKRSKPNMYDK  
 LSRALRYDDKRNIMTKVHKRYAYKDEHNGJAOALQPHPTDESSMYKYSYSDSYMSYH  
 AHQOKVNFVPPHSPMPTSSSFCAASQWYTSPTGIYIPNNVPRHNTHTVSHLS  
 YH

BASE COUNT 669 a 609 c 574 g 542 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,47e-104 Length: 2394  
 Score: 1624.50 Matches: 306  
 Percent Similarity: 79.43% Conservative: 57  
 Best Local Similarity: 66.96% Mismatches: 83  
 Query Match: 66.47% Indels: 11  
 DB: 9 Gaps: 7

US-09-902-772-2 (1-451) x BC010115 (1-2394)

QY 1 MetalaserThrllelysglualeuSerValValSerGluaspGlnSerleupheGlu 20  
 DB 126 ATGAGCGGCGACTATTAAAGAGGCTCTGCGGTGGAGCAGCAGACCACTCCCTTTGAC 185  
 QY 21 CysAlaTyrgly---SerProHisleuAlaTyrglnMetThrlaSerSerSer 39  
 DB 186 TCAGGATGCGAGCGCGCCGACCTATCTCCCAAGGCGGACATGACCTCGGGAGTCT 245  
 QY 40 GluTyrglyGlnThrSerLysMetSerProArGValProGlnGlnAspTrpLeuSerGln 59  
 DB 246 GACATCGGCGAGCCCGCAAGATGCAACCCCTCCACAGCAGAGAGTGGATCAATGAC 305  
 QY 60 ProProAlaArgValThrllelysmetLcYsAsnProAsnGlnValAsnGlySerArg 79  
 DB 306 ---CCAGTGGAGGTCACGTCACAGCGGAGTAT-----GACCACTGAAATGATCCAGG 356  
 QY 80 AsnSerProAspCysSerValAlaAlaTyrglyGlyLysMetValSerSerSerAsp 99  
 DB 357 GAGTCTCCGCTGGAGCTGACGCTTACGAAATGCAAGCTGCGGCGGAGCGAGTCC 416  
 QY 100 ValGlyMetAsnTyrglySerTyrmelGluGlnLysHis---IleProProAsnMet 118  
 DB 417 AACCCCATGAACTTCAACAGCTATATGAGCAGGAAGATGGCCCCCTCCCTCCCAACTG 476  
 QY 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHis 138  
 DB 477 ACCACCAAGAGAGAGATCATGCTCCCGCAGACCCACACTGTCGACACAGAGCAT 536  
 QY 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTyrglyLeuProAspValAspIleLeu 158  
 DB 537 GTGAGGCAATGGCTGGAGTGGCCATTAAGAGATGACAGCTGATGAGATGACACATCC 596  
 QY 159 LeupheGlnAsnIleAspGlyLysGluLeuLysLysMetThrlLysAspAspPheGlnArg 178  
 DB 597 TTTTCCAGACATGATGATGCAAGAACTGTAAATAAGACAAGAGAGACTTCCTCCG 656  
 QY 179 LeuThrProSerTyrglnAlaAspIleLeuLeuSerHisLeuHisTyrglnArgGluArg 198  
 DB 657 GCCACACACCTCTACACACGAGAGTGTGTCACACCTCAGTTACCTCAGGGA--- 713  
 QY 199 GlyAlaThrPheIlePheProAsnThrSerValTyrglnGlnAlaThrGlnArgIleThr 218  
 DB 714 ---AGTTCACGTGGCTTAAATACACCTCCACACCGCAATCTCCACAGATGAGT 770  
 QY 219 ThrArgProAspLeuProTyrglnGlnAlaArgArgSerAlaTrpThrSerHisSerHis 238  
 DB 771 GTCAAGAAAGACCTTCTTATGACTCACTCAAGAGAGAGCTTGGGCAATACATGAT 830  
 QY 239 ProThrGlnSerLysAlaThrGln---ProSerSerSerThrValProLysThrGluasp 257  
 DB 831 TCTGCGCTCAACAAAGTCTCTCCCTTGGAGGGGCACAAACGATCAGTAAGATAACAGAG 890

QY 258 GlnArgProGlnLeuAspProTyrglnIleLeuGlnTyroThrSerSerArgLeuAlaAsn 277  
 DB 891 CACGCGCCCGACCGACGATCCGATACGATCTCGGCGCGACCGACGAGTCCCTACGCCAAC 950  
 QY 278 ProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSer 297  
 DB 951 CCGTGAACCGGCGAGATCCACTGCGGCAATTCCTCCGAGAGCTCTCCGACAGCGCC 1010  
 QY 298 AsnSerAsnCysIleThrTrpGluGlnArgLysSerLysProAsnMetAsnTyrglnSer 317  
 DB 1011 AACGCGACCTGTATCACCCTGGAGGAGGACCAAGGCGGAGTCAAAATACGAGCCGAT 1070  
 QY 318 GluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyrglnSer 337  
 DB 1071 GAGGTGGCGAGCGCTGGGGGAGGAGGAAAGCAAGCCCAACATGAAATACGACAAAGCTG 1130  
 QY 338 SerArgAlaLeuArgTyrglyTyrglyAspLysAsnIleMetThrlLysValHisGlyLysArg 357  
 DB 1131 AGCGGCGCCCGCTTATTACTATGATATAAAACATATAGACCAAGTCCAGCGCAAGA 1190  
 QY 358 TyralaTyrglyPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu 377  
 DB 1191 TATGCTTCAAAATTTGACTTCCACGGCATTTGCCAGGCTCTGACGCCATCCGACGAG 1250  
 QY 378 SerSerMetTyrglyTyrglyProSerAspLeuProTyrmelSerSerTyrglnHisAlaHisPro 397  
 DB 1251 TCGTCATGTACAGTACCTTCTGACATCTCTTACATGCTCTCTTACCATGCCCCACG 1310  
 QY 398 GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhe 417  
 DB 1311 CAGAAAGGGAATTTGCTCCGCCATCCATCCATCCATGCGTCGACATCTCCGACGCTTC 1370  
 QY 418 PheAlaAlaProAsnProTyrglyTrpAsnSerProThrArgLysIleTyrglnProAsn----- 435  
 DB 1371 TTTGGAGCGGATCAATTAATGACCTGACCTCCCGCAGGGGGGAATCTACCCCAACCCCAAC 1430  
 QY 436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlnTyrglyTyrgly 451  
 DB 1431 GTCCCGCCCATCTCAACACCCAGCTGCTTACACTTACAGCAGCTACTAC 1481

Search completed: July 28, 2003, 06:34:10  
 Job time : 3619.89 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## SUMMARIES

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:05:39 ; Search time 289.339 Seconds

(without alignments)  
3510.246 Million cell updates/sec

Title: US-09-902-772-2

Perfect score: 2444

Sequence: 1 MASTIKALSVSEDSLEF.....IYNTRLPAHMPSHLGTYY 451

Scoring table:

BIOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ -p2n.model -DEV-xlh  
-O=/cgn2\_1/USPTO.spool/US09902772/runal\_23072003\_093656\_14868/app\_query.fasta\_1.1294  
-DB-N Geneseq\_101002 -QEMT-fastlap -SUFFIX-rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.ccl  
-LIST=45 -DOCALLIG=200 -THR SCORE-pct -THR MAX=100 -THR MIN=0 -ALIC=15  
-MOB-LOCAT -OUTPRT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09902772 -CGEN\_1\_1\_364 -trunal\_23072003\_093656\_14868 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	2444	100.0	1447	20	AAX26551	DNA encoding chick
2	2414.5	98.8	1528	20	AAX26552	DNA encoding chick
3	1624.5	66.5	2938	14	AAO50644	Human Hum-F11-1 ge
4	1624.5	66.5	2957	24	ABR84139	Human cDNA differe
5	1599	65.4	2954	22	AAH02915	Human shear stress
6	1506.5	61.6	1890	14	AAO50662	Human Hum-F11-1 ge
7	948	38.8	567	22	ABA48124	Human breast cell
8	948	38.8	567	22	ABA66003	Human foetal liver
9	948	38.8	567	22	ABA33090	Probe #11556 for g
10	948	38.8	567	22	AAK14426	Human brain expres
11	948	38.8	567	22	AAK40160	Human bone marrow
12	948	38.8	567	22	AAI20932	Probe #10865 for g
13	948	38.8	567	22	AAI46176	Probe #14862 used
14	948	38.8	567	22	AAI06643	Probe #6634 used t
15	948	38.8	567	24	ABSI4215	Human genome-deriv
16	863	35.3	473	22	ABA50193	Human breast cell
17	863	35.3	473	22	ABA68128	Human foetal liver
18	863	35.3	473	22	ABA35152	Probe #13618 for g
19	863	35.3	473	22	AAK16512	Human brain expres
20	863	35.3	473	22	AAK42265	Human bone marrow
21	863	35.3	473	22	AAI48359	Probe #17025 used
22	863	35.3	473	22	ABR44114	Probe #8687 used t
23	712	29.1	549	24	ABR44114	CDNA #54 encoding
24	654	26.8	533	21	AAOC1342	Human secreted pro
25	650	26.6	420	22	ABA45056	Human breast cell
26	650	26.6	420	22	ABA55529	Human foetal liver
27	650	26.6	420	22	ABA25239	Probe #3705 for ge
28	650	26.6	420	22	AAK03760	Human brain expres
29	650	26.6	420	22	AAK29224	Human bone marrow
30	650	26.6	420	22	AAI35180	Probe #3866 used t
31	650	26.6	420	22	AAI03692	Probe #3683 used t
32	650	26.6	454	22	ABA42995	Human breast cell
33	650	26.6	454	22	ABA53411	Human foetal liver
34	650	26.6	454	22	ABA23187	Probe #1653 for ge
35	650	26.6	454	22	AAK01681	Human brain expres
36	650	26.6	454	22	AAK27129	Human bone marrow
37	650	26.6	454	22	AAI11718	Probe #1651 for ge
38	650	26.6	454	22	AAI33025	Probe #1711 used t
39	650	26.6	454	22	AAI01646	Probe #1637 used t
40	650	26.6	454	24	ABSO1685	Human genome-deriv
41	599.5	24.5	2025	23	ABLO5371	Drosophila melanog
42	562.5	23.0	899	23	ABLI2097	Drosophila melanog
43	511	20.9	1752	22	AAI513672	CDNA encoding Ratt
44	511	20.9	1752	24	AAI38753	Rat lambda73 CDNA
45	498	20.4	318	21	AAI21875	Human secreted pro

## ALIGNMENTS

RESULT 1  
AAX26551  
ID AAX26551 standard; DNA: 1447 BP.

AC AAX26551:

DT 14-JUN-1999 (first entry)

XX DNA encoding chicken C-11 protein.

DE Chicken: C-11 protein; cell calcification inhibiting activity;

XX cell calcification inhibiting agent; c-ery protein; arthritis deformans;

KW ossification; spinal column ligament; ss.

XX Gallus sp.

OS Key Location/Qualifiers  
FH CDS 63..1418

```

FT      /+tag- a
XX      JP11075871-A.
XX      23-MAR-1999.
XX      29-MAY-1998; 98JP-0166076.
XX      20-JUN-1997; 97US-0050297.
XX      18-JUN-1997; 97US-0878177.
XX      (CHUS ) CHUGAI PHARM CO LTD.
XX      (TYPE-) UNIV PENNSYLVANIA.
XX      WPI: 1999-257708/22.
XX      P-PSDB: AAY01520.
XX      An active protein for inhibiting cell calcification - useful for
XX      measuring the calcification of a cell, for diagnosing arthritis
XX      deformans or ossification of spinal column ligament
XX      PS      Disclosure; Page 7-8; 15pp; Japanese.
XX      CC      The present sequence encodes a chicken C-11 protein which has cell
XX      CC      calcification inhibiting activity. The specification also describes
XX      CC      a cell calcification inhibiting agent containing c-ery protein
XX      CC      (AAY01521). The proteins are used for measuring the calcification of a
XX      CC      cell, for diagnosing arthritis deformans or ossification of spinal column
XX      CC      ligament.
XX      SO      Sequence 1447 BP; 440 A; 374 C; 317 G; 316 T; 0 other:

Alignment Scores:
Pred. No.:      2,58e-197      Length:      1447
Score:          2444.00      Matches:      451
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
Gaps:          20              0

US-09-902-772-2 (1-451) x AAX26551 (1-1447)
QY      1 MetLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      63 ATGGCAAGCACTATTAAAGAGCAATATATCATGTGTGATGAAAGCCAGTCTTGTTCAG 122
QY      21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40
DB      123 TGATCCCTACGGATGCGCCCACTTGCAAGACAGAAATGACAGCCTCTTCCAGTGA 182
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
DB      183 TATGGCGAAACATCAAGATGAGAGCCCGCGGTGCCAGCAGAGCTGTTATCACAGCCC 242
QY      61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerAsn 80
DB      243 CCGGCCGAGTATACCATTAAGATGAGTGAACCAACCAAGTAAATGGTCAAGGAAT 302
QY      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB      303 TCACCTATGATCTGACACGCGGCAAAAGAGGAAATGGTTAGCAATTCAGACAATGTT 362
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThr 120
DB      363 GGGATGAACATATGGAACCTACATGAGAGAGAGATATTCGGCTCCAAATATGACAAACC 422
QY      121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140
DB      423 AATGAAAGAGAGATTATTTGCCAGCAGATCTTACGTTATGAGACAGACCAATGTCGG 482
QY      141 GlnTyrLeuGlnTyrProAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
DB      483 CAGTGGCTGAGTGGGAGTGAAGAGATATGTTCTTCAGACGTTGACATCTGTGTTTC 542

```

```

QY      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB      543 CAGAACATGTGATGGGAAAGAGTTGTGTAAATATGACCAAGATGATCTCCAGAGACTCAGC 602
QY      181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAla 200
DB      603 CCGAGCTATACGAGATATCTCCCTGTCACACTACCTACCTCAGAGAGAGAGAGGCC 662
QY      201 ThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArg 220
DB      663 ACTTTTATTTTCCAAATATCATCATGTTTACCAGAGACCAACGCAAGAAATTAACACAGG 722
QY      221 ProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrThrSerHisSerHisProThr 240
DB      723 CCAGATTTTACCTTATGAGCAAGCAGAGATACAGCTGAGAGATCAGACCTCACCTCCACT 782
QY      241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgPro 260
DB      783 CAGTCAAAAGCTTACCCCAACATCATCTTCAACAGTGCACAAACAGAAAGACAGCCTGCT 842
QY      261 GlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
DB      843 CAGTTATATCTTATCATGATTTCTTGACCGACAGCAGCCGCTGTGCAATTCAGGAGT 902
QY      281 GlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsn 300
DB      903 GGGCAGATACAGTATGAGAGTTCCTACTGAGAGCTTCTGTCGACAGCTCCAACTCCAAC 962
QY      301 CysIleThrThrProGluGlnTyrAsnGlyGluPheLysMetThrAspProAspGluValAla 320
DB      963 TGCATCACCTGGGAGGCGCAAAATGGGAGTTCAGATGACAGACCTGTGTAAGTGGCT 1022
QY      321 ArgArgTyrProGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAla 340
DB      1023 CCGCGTTGGGAGAGAGAGAAAGCAACCTAATCATGATGACAAACTCAGCCGTGCA 1082
QY      341 LeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360
DB      1083 CTTGCGTACTACTATGACAAAAATATATGACTTAAGTTTCATGTTAAGCTTATGCTTAC 1142
QY      361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerSerMet 380
DB      1143 AAATTTATTTCCAGGAATGCTCAGGCCCTCCAGCTTCACTCAGATATATCATG 1202
QY      381 TyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMet 400
DB      1203 TACAAATATCCCATCAGACCTCCCTACATGATGATTCATGATGACACCCAGAGATG 1262
QY      401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAlaAla 420
DB      1263 AACTTGTACCTCCCAATCCCTGCTTCCCGTAACTTCATCCACACTTTTGGCTGCC 1322
QY      421 ProAsnProTyrTyrPAsnSerProThrGlyGlyIleTyrProAsnThrArgLeuProAla 440
DB      1323 CCTATCATACACTGGAATTCACCACTGGAGGATCTTACCCCATACAGGAGCTGCCAGCT 1382
QY      441 AlaHisMetProSerHisLeuGlyThrTyrTyr 451
DB      1383 GCTCATATGCTTCCCATCTTGGCACTACTATC 1415

RESULT 2
AAX26552
ID      AAX26552 standard; DNA: 1528 BP.
XX      AAX26552;
AC      AAX26552;
XX      14-JUN-1999 (first entry)
XX      DNA encoding chicken c-ery protein.
DE      Chicken; C-11 protein; cell calcification inhibiting activity;
KW      cell calcification inhibiting agent; c-ery protein; arthritis deformans;

```

KM ossification; spinal column ligament; ss.  
 XX Gallus sp.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 63..1499  
 FT CDS /\*tag= a  
 JPL1075871-A.  
 PN  
 XX  
 PD 23-MAR-1999.  
 XX  
 PF 29-MAY-1998; 98JP-0166076.  
 XX  
 PR 20-JUN-1997; 97US-0050297.  
 PR 18-JUN-1997; 97US-0878177.  
 XX  
 PA (CHUGAI PHARM CO LTD.  
 PA (UYPE-) UNIV PENNSYLVANIA.  
 DR  
 DR WPI: 1999-257708/22.  
 DR P-PSDB: AA01521.  
 XX  
 PT An active protein for inhibiting cell calcification - useful for  
 PT measuring the calcification of a cell, for diagnosing arthritis  
 PT deformans or ossification of spinal column ligament  
 XX  
 PS Disclosure: Page 8-9; 15pp; Japanese.  
 XX  
 CC The present sequence encodes a chicken c-ery protein. The specification  
 CC also describes a chicken C-11 protein (AA01520) which has cell  
 CC calcification inhibiting activity and a cell calcification inhibiting  
 CC agent containing c-ery protein. The proteins are used for measuring the  
 CC calcification of a cell, for diagnosing arthritis deformans or  
 CC ossification of spinal column ligament.  
 XX  
 SQ Sequence 1528 BP; 464 A; 395 C; 330 G; 339 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8 71e-195 Length: 1528  
 Score: 2414.50 Matches: 450  
 Percent Similarity: 94.14% Conservative: 0  
 Best Local Similarity: 94.14% Mismatches: 1  
 Query Match: 98.79% Indels: 27  
 DB: 20 Gaps: 1  
 US-09-902-772-2 (1-451) x AA026552 (1-1528)  
 QY 1 MetalaSerThrIleuysGluAlaLeuSerValSerGluAspGlnSerLeuPheGlu 20  
 DB 63 ATGGCAACGACTATTAAAGAAAGCATTTATCAGTGGTGAAGTGAAGACCACTCTTGTGTAG 122  
 QY 21 CysalaTyrglySerProHisIleuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 DB 123 TGTGCTACGAGATGCCACCCTTGCAAAAGACAAATGACAGCCTCTCTTCCAGTAA 182  
 QY 41 TyrGlyGlnThrSerIleuMetSerProArgValProGlnGlnAspTIPLeuSerGlnPro 60  
 DB 183 TATGGGCAAAACATCAAAATGAGCCCGCGCTCCCGCAGCAGACTGGTTATCAACAGCCC 242  
 QY 61 ProAlaArgValThrIleuysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 DB 243 CCGGCGACAGTATACCATTAATGAGAGTGAATCCCAACGATTAAATGGGCAAGGAT 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyIleuysMetValSerSerSerAspAsnVal 100  
 DB 303 TCACCTGATGACGTGACGCGTGGCAAAAGAGGAAATAGTTAGACAGTTACAGACAATGTT 362  
 QY 101 GlyMetAsnTyrglySerTyrMetGlnGlnLysHisIleProProProAsnMetThrThr 120  
 DB 363 GGGATGAATATGGAAGCTACATGGAAGAAAGCATTAATCCGCTCCAAATATGCAACCC 422  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140

DB 423 AATGAACGAGAGTATTGTGCGACAGATCTCTAGCTTATGGAGCAGACCATGATACCG 482  
 QY 141 GlnTPLeuGluTTPAlaValLysGluTyrglyLeuProAspValAspIleLeuLeuPhe 160  
 DB 483 CAGTGGCTGGAGTGGGACAGTGAAGAGTATGCTTCCAGACGTGACATCTGTGTGTTTC 542  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 DB 543 CAGAACATTTGATGGAAAGAGTTGTGTAATAAGCAAAAGATGACTTCCAGAGACTCAGC 602  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluArg----- 198  
 DB 603 CCGACCTTAACAGCAGATATCTCTCTGCACACCTTACACTTACCTAGAGAGACTCTCTT 662  
 QY 198 ----- 198  
 DB 663 CCACATTTGACTTACGATGATGTGATGAAGCCCTTACAAAACTCTCCAGGTTAATGCAT 722  
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213  
 DB 723 GCTAGAAACACAGAGAGGCCACTTTTATTTTCCAAATACATCAGTTTACCCAGAGACA 782  
 QY 214 ThrGlnArgIleThrThrArgProAspLeuProTyrglnGlnAlaArgArgSerAlaTyrP 233  
 DB 783 ACGCAAGAAATACACAAAGGACAGATTTACTTATGAGCAAGCAGAGGATGACAGCTGG 842  
 QY 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValPro 253  
 DB 843 ACGAGTCAACAGCAGTCCACATCAGTCAAAAGTACCAACATCATCTTCAACAGTGGCC 902  
 QY 254 LysThrGlnAspGlnArgProGlnLeuAspProTyrglnIleuGlyProThrSerSer 273  
 DB 903 AAAACAGAGACACAGCTCTCAGTTAATCCTTATCAGATTTCTGGAACCGCAGCAGAC 962  
 QY 274 ArgLeuAlaAsnProGlnSerGlyGlnIleGlnLeuTyrPGLnPheLeuLeuLeuLeu 293  
 DB 963 CGCTTTCGAATCCAGGAGAGTGGCAGATACAGCATGAGCTTCTCTGAGACTTCTG 1022  
 QY 294 SerAspSerSerAsnSerAsnGlyIleThrTPGlnGlyTyrThrAsnGlyLubPheLysMet 313  
 DB 1023 TCGGACACTCCAACTCCAACTGATCAGCTGGAGGAGCAAAATGGGAGATTCAAGATG 1082  
 QY 314 ThrAspProAspGlnValAlaArgArgTyrPGLnGluTyrglySerLysProAsnMetAsn 333  
 DB 1083 ACAGACCTGATGAAGTGGCTCGGCTGGGAGAGAGAGAAAGCAAACTTAACATGAAAC 1142  
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleuMetThrLysVal 353  
 DB 1143 TATGCAAACTCAGCCGCTGCACCTTCTACTACTATGACAAATAATATATGACTAAAGTT 1202  
 QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373  
 DB 1203 CATGTTAAACCTGATGCTTACAAATTTATTTCCACGGAATGCGTCAGGCCCTCAGGCT 1262  
 QY 374 HisProProGluSerSerMetLysTyrTyrProSerAspLeuProTyrglyMetSerSerTyr 393  
 DB 1263 CACCTCCAGATATCATCATGACAAATACCATCAGACCTCCCTCATCATGATGTTCTTAC 1322  
 QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisIleProProAlaLeuProValThr 413  
 DB 1323 CATGCAACCCCGCAAAATATGATTAATTTAGCTCCCAATCCCGCTTTCGCCGTAAAC 1382  
 QY 414 SerSerSerPhePheAlaAlaProAsnProTyrrTPAsnSerProThrGlyIleTyr 433  
 DB 1383 TCATCCAGCTTTTGTGCGCCCTAATATCATGAGATTCACCAACTGAGGAGCATTTAC 1442  
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 DB 1443 CCCAATACAGAGTGGCAGCTGCTCATATATGCTTCCCATCTTGGACACTTACTAC 1496  
 RESULT 3  
 AA050644

ID AA050644 standard; cDNA: 2938 BP.  
 XX AA050644;  
 AC  
 XX  
 XX 26-MAY-1994 (first entry)  
 DT  
 XX  
 DE Human Hum-Fli-1 gene clone BM025.  
 XX  
 KW chromosomal translocation; chimeric; chimeric; Ewing sarcoma;  
 KW Ews gene; malignant melanoma; hum-fli-1;  
 KW primitive peripheral neuroectodermal tumour; human chromosome 11;  
 KW human chromosome 22; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 143..1501  
 FT /\*tag= a  
 FT /product= HUM-Fli-1  
 FT polyA\_signal 2908..2913  
 FT /\*tag= b  
 FT  
 XX MOJ323549-A.  
 XX  
 XX 25-NOV-1993.  
 PD  
 XX 19-MAY-1993; 93WO-FR00494.  
 PF  
 XX 20-MAY-1992; 92FR-0006123.  
 PR  
 XX (CNRS ) CNRS CENT NAT RECH SCI.  
 PA  
 XX Aurias A, Delattre O, Desmaze C, Melot T, Peter M,  
 PI Piongestel B, Thomas G, Zucman J;  
 P1  
 XX WPI: 1993-386580/48.  
 DR P-PSDB: AAR44556.  
 XX  
 XX New nucleic acid of EWS gene and its hybrid(s) - contg. gene  
 PT sequence involved in chromosomal trans-location, also derived  
 PT mRNA, probes, fusion proteins etc., for diagnosis and treatment  
 PT of Ewing sarcoma and melanoma  
 PT  
 XX  
 PS Disclosure: Fig 7; 123pp; French.  
 XX  
 CC The probe 11R1 was used to screen a human marrow cDNA library  
 CC (Clontech cat.# HL1058). The clone BM025 was identified and  
 CC sequenced. It represents the entire coding region together with  
 CC 5'- and 3'-UTRs of the Hum-Fli-1 gene.  
 CC  
 XX  
 SQ Sequence 2938 BP; 847 A; 692 C; 675 G; 724 T; 0 other;  
 SQ  
 Alignment Scores:  
 Score: 1.52e-127 Length: 2938  
 1624.50 Matches: 306  
 Percent Similarity: 79.43% Conservative: 57  
 Best Local Similarity: 66.96% Mismatches: 83  
 Query Match: 66.47% Indels: 11  
 DB: 14 Gaps: 7  
 US-09-902-772-2 (1-451) x AA050644 (1-2938)  
 QY 1 MetaLaserThrlleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGln 20  
 DB 143 ATGAGCGGAGCTATTAGAGGCGCTCTGCTGGTGGAGCGAGCCAGTCCCTCTTGAC 202  
 QY 21 CysAlaTyrGly---SerProHisLeuAlaLysThrLumethrAlaSerSerSer 39  
 DB 203 TCAGCGTAGCGAGCGGAGCGCCATCTCCCAAGGCGACATGAGCTCGCGGAGTCC 262  
 QY 40 GlnTyrGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59  
 DB 263 GACTACGGGAGCGCCCAAGATCAACCCCTCCACACGAGGAGTGGATCAATACG 322

QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79  
 DB 323 ---CCAGTAGAGGCTCAACGTCACACGGGAGTAT-----GACCACATGAATGATCCAGG 373  
 QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99  
 DB 374 GAGTCTCCGTGGTGGTCTGATGAGGATGCAAAATGACAGCAAGTGGCGGAGCGGATCC 433  
 QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGlnLysHis---IleProProAsnMet 118  
 DB 434 AACCCCATGAACCTAACAGCTATATGACAGAGAAGATGGCCCCCTCTCCACATG 493  
 QY 119 ThrThrAsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHis 138  
 DB 494 ACCACCAACAGAGAGAGTATCTCCCGGACAGCCACACACTGTGACACAGAGACAT 553  
 QY 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeu 158  
 DB 554 GTGAGGCAATGGCTGGAGTGGCCATTAAGAGATATAGCTTGATGAGATCGACATCC 613  
 QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178  
 DB 614 TTTTCCAGAAACATGATGGCAAGAACTGTATAATGAACAAGAGACTTCCTCCG 673  
 QY 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgLysArg 198  
 DB 674 GCCACACACCTCTACACAGCAAGAGTCTGTGCACACCTCACTACCTACAGGAA--- 730  
 QY 199 GlyAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThr 218  
 DB 731 ---AGTTCAGTGGCCCTTATATACACTCCACCCAGCAACATCTCACTCACTAGT 787  
 QY 219 ThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThrSerHisSerHis 238  
 DB 788 GTCAAGAGAACCTCTTATGACTAGTCAAGAGAGAGAGTGGGCAATACATGAT 847  
 QY 239 ProThrInseryAlaThrGln---ProSerSerSerThrValProLysThrGlnAsp 257  
 DB 848 TCTGGCTCAACAAATCTCTCCCTTGGAGGGGCAACAGACATAGAAATACAGAG 907  
 QY 258 GlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 277  
 DB 908 CAACGGCCCAACCGAGATCCGTATCAGATCTGGGCCGACACACATCGCTAGCCAC 967  
 QY 278 ProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGlnLeuLeuSerAspSerSer 297  
 DB 968 CTGTGAGCGGAGAGATCCAGCTGTGCAATCTCTCTGAGCTGCTCCGACAGCGCC 1027  
 QY 298 AsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLysMetThrAspProAsp 317  
 DB 1028 AACGCCAGCTGTATCAGCTGGAGGGGACACAGCGGAGGTCAAAATGAGCGACCCGAT 1087  
 QY 318 GluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeu 337  
 DB 1088 GAGGTGGCCAGCGGCTGGGCGAGCGGAAAGACGCCAACAATGAATTAACGAAAGCTG 1147  
 QY 338 SerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArg 357  
 DB 1148 AGCCGGGCCCTCTCTTATCTATGATGAATAAACAATATGACCAAGAGCGACGCAAAAGA 1207  
 QY 358 TyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGln 377  
 DB 1208 TATGCTTACAAATTTGACTTCCACGCGATGCCAGCTGTGACGACACATCCGACGAG 1267  
 QY 378 SerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisPro 397  
 DB 1268 TCGTCATGTACAAAGTACCTTGTGACATCTCTACATGCTCTTCAACATGCGCCACAG 1327  
 QY 398 GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhe 417  
 DB 1328 CAGAAGGTGAACTTTGTCCCTCCGACATCTCTCAATGCTGTCTCACTTCTCCAGCTTC 1387

Oy	418	PhalAlaAlProAsnProTyrTrpAsnSerPcOfTrnGlyGlyIleTyrProAsn-----	435
			:::
Db	1388	TTTGAGACCGCATCAACATATCTGAGACTTCCCCACGGGGGAAATCTACCCACCCAC	1444
Oy	436	---ThrArgLeuProAlaIahIshMetProSerHisLeuGlyTrhTyrTyr 451	
Db	1448	GTCCCGCGCATCTCAACACACCGAGCGCTTCACACTTAGCAGACTACTAC	1498
RESULT 4			
ABK84139			
ID	ABK84139	standard; cDNA; 2957 BP.	
XX	AC	ABK84139;	
XX	DT	14-AUG-2002 (first entry)	
DE	XX	Human cDNA differentially expressed in granulocytic cells #710.	
XX	XX		
KW	KW	Human; ss; granulocytic cell; DNA chip; bacterial infection;	
KW	KW	viral infection; parasitic infection; protozoal infection;	
KW	KW	fungal infection; sterile inflammatory disease; psoriasis;	
KW	KW	rhematoid arthritis; glomerulonephritis; asthma; thrombosis;	
KW	KW	cardiac reperfusion injury; renal reperfusion injury; ARDS;	
KW	KW	adult respiratory distress syndrome; inflammatory bowel disease;	
KW	KW	Crohn's disease; ulcerative colitis; periodontal disease;	
XX	KW	granulocyte activation; chronic inflammation; allergy.	
OS	XX	Homo sapiens.	
XX	XX		
PN	XX	W0200228999-A2.	
PD	XX	11-APR-2002.	
XX	XX		
PF	XX	03-OCT-2001; 2001WO-US30821.	
XX	XX		
PR	XX	03-OCT-2000; 2000US-237189P.	
XX	XX		
PA	XX	(GENE-) GENE LOGIC INC.	
PI	XX	Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;	
DR	XX	WPI; 2002-435328/46.	
XX	XX		
PT	XX	Detecting granulocyte activation by detecting differential expression	
PT	XX	of genes associated with granulocyte activation, which serves as	
PT	XX	diagnostic markers that is useful for monitoring disease states and	
PT	XX	drug toxicity -	
XX	XX		
PS	XX	Claim 1; SEQ ID NO 710; 114pp; English.	
XX	XX		
CC	CC	The invention relates to detecting (M1) granulocyte (GC) activation	
CC	CC	(GCA), by detecting the level of expression of gene(s) (Gs) identified by	
CC	CC	DNA chip analysis as given in the specification, and comparing	
CC	CC	the expression level to an expression level in an unactivated	
CC	CC	GC, where differential expression of Gs is indicative of GCA.	
CC	CC	Also included are modulating (M2) GA by contacting GC with an agent	
CC	CC	that alters the expression of at least one gene in Gs; (2) screening (M3)	
CC	CC	for an agent capable of modulating GCA or an inflammation (especially	
CC	CC	chronic) in a tissue, an allergic response in a subject, exposure of a	
CC	CC	subject to a pathogen or sterile inflammatory disease using the	
CC	CC	gene expression profile; (3) detecting (M4) an inflammation (especially	
CC	CC	chronic) in a tissue, an allergic response in a subject, exposure of a	
CC	CC	subject to a pathogen or sterile inflammatory disease, by detecting the	
CC	CC	level of expression in a sample of the tissue of gene(s) from Gs, where	
CC	CC	the level of expression of the gene is indicative of inflammation;	
CC	CC	(4) treating (M5) an inflammation (especially chronic) or in a tissue,	
CC	CC	an allergic response in a subject, exposure of a subject to a pathogen	
CC	CC	or sterile inflammatory disease, by contacting a tissue having	
CC	CC	inflammation with an agent that modulates the expression of gene(s)	
CC	CC	from Gs in the tissue. M1 is useful for detecting GCA. M2 is useful for	
CC	CC	modulating GA; M3 is useful for screening an agent capable of modulating	
CC	CC	GCA preferably in an inflammation in a tissue; M4 is useful for	
CC	CC	detecting an inflammation (especially chronic) in a tissue, an allergic	

	CC	response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
	CC	glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
	CC	reperfusion injury, ARDS, adult respiratory distress syndrome,
	CC	inflammatory bowel disease, Crohn's disease, ulcerative colitis,
	CC	periodontal disease; also bacterial infection, viral infection,
	CC	parasitic infection, protozoal infection, fungal infection and M5 is
	CC	useful for treating one of the above conditions. The present
	CC	sequence represents a gene differentially expressed in granulocytes.
	CC	Note: The sequence data for this patent did not form part
	CC	of the printed specification, but was obtained in electronic
	CC	format directly from WIPO at
	CC	ftp.wipo.int/pub/published_pcl_sequences.
	xx	
SQ	Sequence	2957 bp; 855 A; 698 C; 680 G; 724 T; 0 other;
	Alignment Scores:	
Pred. No.:	1.53e-127	Length: 2957
Score:	1624.50	Matches: 306
Percent Similarity:	79.43%	Conservative: 57
Best Local Similarity:	66.96%	Mismatches: 83
Query Match:	66.47%	Indels: 11
DB:	24	Gaps: 7
US-09-902-772-2 (1-451) x ABR84139 (1-2957)		
OY	1	MetAlasSerThrIleLysGluAlaLeuSerValIserGeluspInserLeuPheGln 20
Db	173	ATGGACGGAGCATTTTAAAGAGGCTGTGTGGTGTAAGCAGCACCAAGTCCTTTTGAC 232
OY	21	CysAlaTryGly--SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
Db	233	TGAGCGTACGGAGCGGCAGCCATCTCCCCAAGGCCGACATGATGACTCTTCGGGAGTCT 292
OY	40	GluTryGlyGlnThrSerLysMetSerProArgValProGlnInsprTrpleLysSerGln 59
Db	233	GACTACGGGACACCCCACAGATCAACCCCTCCACCACAGGAGGATGATCAATCAG 352
OY	60	ProProlaArgValThrIleLysMetGluCysAsnPrpAsnGlnAlaSnngLysSerArg 79
Db	353	--CCAGTGAGGGGTCAAGCTCAAGCGGAGTAT-----GACCACATGAATGATCCAGG 403
OY	80	AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db	404	GAGTCTCCGGTGGACTGTGACGGCTTGACAATCAGCAACCTGGTGGCGGAGGCGAGTCC 463
OY	100	ValGlyMetAsnTryGlySerTryMetGluGluLysHis--IleProProAsnMet 118
Db	464	AACCCCATGAATTACAAACGCTATATGACGAGAGAATAAGCCCCCTCTCCCAACATG 523
OY	119	ThrThrAsnGluArgArgValIleValProAlaAspProThrIleUrpSerThrAspHis 138
Db	524	ACCACCAACGAGAGAGATCATCGCCCCCGAGACCCACACTGTGGACACAGGAGCAT 583
OY	139	ValArgGlnTrpLeuGlnThrPraLalaLysGluTryGlyLeuProAspPraLaspIleLeu 158
Db	584	GTAGGCGCATGGCTGGAGTGGCCATAAAGAGTACAGCTTGATGGAGATCGCACATCC 643
OY	159	LeuPheGlnAsnIleAspOlyLysGluLeuCyLSysMetThrLysAspAspPheGlnArg 178
Db	644	TTTTTCCAGAACATGATGGACGAAGCAAGCTGTAAATGAACAAGAGAGGACTCTCCGCG 703
OY	179	LeuthrProSerTryArgAsnAlaSpIleLeuLeuSerHisLeuHisLysTrileuArgGluArg 198
Db	704	GCCACCAACCTCTTACAAACAGGAAGTGGCTGTGTGCACACCTCAAGTTAACTCAGGAA--- 760
OY	199	GlyAlaThrPheIlePheProAsnThrSerValTryProGlnAlaIleThrGlnArgIleThr 218
Db	761	---ACTTCACGCTGGCCTATATACAAACCTCCACACCGACCAATCCCTCAGATTGACT 817
OY	219	ThraArgProAspLeuProTryGluGlnAlaArgArgSerAlaTrpThrSerHisSerHis 238
Db	818	GTCAAAGAGACCCCTTTATGACATCAGTCAAGAGAGGAGCTTGGGGCAATTAACATGAAT 877

```

QY 239 ProthrGlnSerLysAlaThrGln---ProSerSerThrValProLysThrGluAsp 257
Db 878 TCTGGCTCAACAAAGTCCCTCCCTGGAGGGCACAAAGATCATGATACACAG 937
QY 258 GlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 277
Db 938 CAACGGCCCGCCAGCCAGATCCGTATCATGATCCTGGGGCCGACACAGATCGCTAGCCAAAC 997
QY 278 ProGlySerGlyGlnIleGlnLeuThrGlnPheLeuGlnIleLeuSerSerSerSer 297
Db 998 CTTGGAAGCGGCGATCATGCTGCTGCAATTCCTCTGGAGCTGCTCTCCGACAGCGCC 1057
QY 298 AsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyIlePheLysMetThrAspProAsp 317
Db 1058 AACCCAGCTGTATCATCTGGAGGGAGCCAGCGGGAGTTCAAAATAGACGACCCCGAT 1117
QY 318 GluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeu 337
Db 1118 GAGGTGGCCAGGCGCTGGGGGAGCGAAGCAAGCCCAATGAAATTCAGACACCTG 1177
QY 338 SerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValIleGlyLysArg 357
Db 1178 AGCCGGGCGCTCCCTTATCTATGATPAAACATTTATGACCAAGTGCACGCGAAGA 1237
QY 358 TyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu 377
Db 1238 TATGCTTACAAATTTGACTTCCAGCGCATTCGCCAGCTCGACGCCAATCCACAGCAG 1297
QY 378 SerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisPro 397
Db 1298 TCGTCACATGTACAGTACCTCTTCGACATCTCTACATGCTCTCTACATGCCACAG 1357
QY 398 GlnLysMetAsnPheValAlaProHisProProAlaLeuProValIleThrSerSerPhe 417
Db 1358 CAGAGGTGAATCTTGTCCCTCCCATCCATCCATCCATGCTGCTGCTCTCTCCAGCTTC 1417
QY 418 PheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyIleTyrProAsn----- 435
Db 1418 TTTCGACCGCATATCATATCTGACCTCCGCCAGGGGGGAAATCTACCCCAACCCCAAC 1477
QY 436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db 1478 GTCCCGCCGATCCTTAACACCCAGCGCTTCACACTTAGGACACTTACTAC 1528

RESULT 5
AAH02915
ID AAH02915 standard; DNA; 2954 BP.
AC AAH02915;
DT 15-JUN-2001 (first entry)
XX
DE Human shear stress-response coding sequence SFO ID NO: 83.
XX
KW Human; shear stress-response protein; vascular disease;
XX
OS Homo sapiens.
XX
PN WO200125427-A1.
PD 12-APR-2001.
PF 02-OCT-2000; 2000WO-JP06840.
PR 01-OCT-1999; 99JP-0280976.
XX
PA (KYOWA) KYOWA HAKKO KOYO KK.
XX
PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
Kuga T, Sekine S, Nakamura Y, Sugano S;

```

```

XX WPI: 2001-266308/27.
DR P-PSDB: AAB90792.
XX
PT DNA sequences, proteins encoded by them and antibodies against them
PT useful in diagnosis and treatment of vascular disease caused by
PT arteriosclerosis -
XX
PS Claim 20; Page 462-466; 678pp; Japanese.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension.
XX
SQ Sequence 2954 BP; 846 A; 698 C; 683 G; 727 T; 0 other;

Alignment Scores:
Pred. No.: 2,2e-125 Length: 2954
Score: 1599.00 Matches: 304
Percent Similarity: 78.99% Conservative: 57
Best Local Similarity: 66.52% Mismatches: 84
Query Match: 65.43% Indels: 12
DB: Gaps: 8

US-09-902-772-2 (1-451) x AAH02915 (1-2954)
QY 1 MetAlaSerThrIleLysGluAlaLeuSerValIleSerGluAspGlnSerLeuPheGln 20
Db 173 ATGAGCGGAGCTATTAGAGAGGCTCTGCTGGTGAGCAGCAGCCAGTCCCTTGTGAC 232
QY 21 CysAlaTyrGly---SerProHisLeuAlaLysThrGlnMetThrAlaSerSerSer 39
Db 233 TCAGCTGACGAGCGGAGCGGCCATCTCCCAAGCGCGCATGATGATGCTCTGGGAGTCT 292
QY 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59
Db 293 GACTAGGCGGAGCGGCCCAAGATCAACCCCTCCACACAGCAGAGTGATCAATCAG 352
QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db 353 ---CCAAGTGGGGTCAACGTCACGCGGAGATAT-----GACCACATGAATGATCCAG 403
QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db 404 GAGTCTCGGTGGAGTCTACAGCTTAGCAAAATGCAGCAAGCTGGTGGGCGAGCGAGTCC 463
QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---IleProProAsnMet 118
Db 464 AACCCCATGAATCAACACAGCTATATGAGAGAGAAATGGCCCCCTCTCCCAACATG 523
QY 119 ThrThrAsnGluArgValIleValProAlaAspProThrIleuTrpSerThrAspHis 138
Db 524 ACCACCAACGAGAGAGATCATCTCCCGCAGACCCCACTGTGTGACACAGAGCAT 583
QY 139 ValArgGlnTrpLeuGlnTrpAlaValLysGlyTyrGlyLeuProAspValIleLeu 158
Db 584 GTGAGCAATGGCTGAGGTGGCCATTAAGAGATACGTTGATGGAGATGCACATCC 643
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
Db 644 TTTTTCAGAAATGATGTCAGAGAACTGTAAATGAAACAAGAGGACTTCCTCCCGC 703
QY 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg 198
Db 704 GCCACCAACCTCTAACAACAGAGAGTCTGTGTCACTCACTCACTTACCTCAGAGAA--- 760
QY 199 GluAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThr 218
Db 761 ---AGTTCAGTGGCTTATATCAACCTCCACACCGCAATTCCTCAATTTGAGT 817
QY 219 ThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThrSerHisSerHis 238

```







DE Human breast cell single exon nucleic acid probe #6819.  
 XX Human: microarray: single exon probe; gene expression; breast;  
 KW disease; cancer; ss.  
 XX Homo sapiens.  
 OS  
 PN MO200157271-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-0500662.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-496933/54.  
 XX  
 PT New spatially-addressable set of single exon nucleic acid probes,  
 PT useful for measuring gene expression in sample derived from human  
 PT breast, comprises number of single exon nucleic acid probes  
 XX  
 PS Claim 4; SEQ ID NO 6819; 327pp + sequence listing; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting  
 CC the probes with a collection of detectably labelled nucleic acids  
 CC derived from mRNA of human breast, and then measuring the label  
 CC bound to each probe of the microarray. The probes are useful for  
 CC verifying the expression of regions of genomic DNA predicted to  
 CC encode proteins. They are useful for gene discovery, and for  
 CC determining predisposition and/or prognosing breast disease. Gene  
 CC expression analysis is useful for assessing the toxicity of chemical  
 CC agents on cells. The microarray of this invention presents a far greater  
 CC diversity of probes for measuring gene expression, with far less bias  
 CC than expressed sequence tag microarrays. The method is suitable for  
 CC rapid production of functional information from genomic sequence. The  
 CC present sequence is a single exon nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2,91e-71 Length: 567  
 Score: 948.00 Matches: 169  
 Percent Similarity: 97.74% Conservative: 4  
 Best Local Similarity: 95.48% Mismatches: 4  
 Query Match: 38.79% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-902-772-2 (1-451) x ABA48124 (1-567)  
 QY 275 LeuAlaAsnProGlySerGlyInIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294  
 DB 534 CTGACCTCATTAAGCGGCGGCGAGATCCAGCTTTGGAGTTCCTCTGGAGCTCGTGC 475  
 QY 295 AspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlnGluPheLysMetThr 314  
 DB 474 GAAAGCTCCACAGTCGAGCTGACCTGGGGAAGGACCAACGCGGAGTTCAAGATGACG 415  
 QY 315 AspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyr 334

DB 414 GATCCGACAGAGGCGCCGCGCTGGGAGAGCGGAAGCAACCAACATGAAGTAC 355  
 QY 335 AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 354  
 DB 354 GATAGGCTCAGCCGCGCCCTCCGTACTACTATGACAAGAACATCATGACCAAGGTCCAT 295  
 QY 355 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374  
 DB 294 GGAAGGCTAGCGCTACAAAGTTGAGTCTCAGGGAGGCGCCGCTCCAGCCCCAC 235  
 QY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerTyrHis 394  
 DB 234 CCCCCGAGTCACTCTGTACAGTACCCCTCAGACCTCCGTAACATGGGCTCCTATCAC 175  
 QY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414  
 DB 174 GCCCACCACAGAGAGAACTTGTGGCCCCCACCCTCCAGCCCTCCGTCGACATCT 115  
 QY 415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleTyrPro 434  
 DB 114 TCCAGTTTTTTTGTGCTGCCCAAAACCCATCTGGAATTCACCAACTGGGGTTATATACCC 55  
 QY 435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 DB 54 AACACTAGCTCCGCCACCAACCAATATGCCCTTCATCTGCGCACTTACTAC 4  
 RESULT 8  
 ABA66003/c  
 ID ABA66003 standard; DNA; 567 BP.  
 XX  
 AC ABA66003;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Human foetal liver single exon nucleic acid probe #14308.  
 XX  
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200157271-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-0500669.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-483447/52.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human foetal liver -  
 XX  
 PS Claim 4; SEQ ID NO 14308; 639pp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC foetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other:

## Alignment Scores:

Pred. No.: 2,91e-71 Length: 567  
 Score: 948.00 Matches: 169  
 Percent Similarity: 97.74% Conservative: 4  
 Best Local Similarity: 95.48% Mismatches: 4  
 Query Match: 38.79% Indels: 0  
 DB: 22 Gaps: 0

US-09-902-772-2 (1-451) x ABA66003 (1-567)

QY 275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294  
 DB 534 CTGACCTCATTTAGGACAGTGGCCAGATCCAGCTTGGGAGTTCTCTCGAGCTCTGCG 475  
 QY 295 AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphelusMetThr 314  
 DB 474 GACAGCTCCAACTCCAGCTCATCAGCTGGGAGGCGCAACGGGAGTTCAAGATGACG 415  
 QY 315 AspProAspGluValAlaIleArgTrpGlyGluArgLysSerLysProAsnMetAsnTyr 334  
 DB 414 GATCCCGACAGAGTGGCCCGCTGGGAGAGCGGAGGCAAAACCAACATGAACTAC 355  
 QY 335 AspLysLeuSerArgAlaIleArgTrpGlyTyrAspLysAsnIleMetThrLysValHis 354  
 DB 354 GATAGCTGACGCGCGCCCTCCCTGACTATGACAAAGAACATCATGACCAAGTTCAT 295  
 QY 355 GlyLysArgTrpAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374  
 DB 294 GGGAGAGCGCTACCGCTCAAGTTGCACTTCCAGGGGATCCCGACGGCTCCAGCCGAC 235  
 QY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394  
 DB 234 CCCCCGAGTCACTCTGTACAGTACCCCTCAGACCTCCCGTACATGGGCTCTATCAC 175  
 QY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414  
 DB 174 GCCACCCACAGAAAGTGAATCTTGTGGGCCCCACCCCTCCAGCCCGGAGACATCT 115  
 QY 415 SerSerPhePheAlaAlaProAsnProTyrTTPAsnSerProThrGlyIleTyrPro 434  
 DB 114 TCCAGTTTCTTCTGCTGCCCAACCCATCTGGAATTCACCACTGGGGGTATATACCC 55  
 QY 435 AspThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 DB 54 AACACTAGGCTCCCAACGACCATATGCTTCTCATCTGGGCACTTACTAC 4

## RESULT 9.

ABA33090/c  
 ID ABA33090 standard; DNA: 567 BP.

XX ABA33090;

XX 23-JAN-2002 (first entry)

DE Probe #11556 for gene expression analysis in human heart cell sample.

XX Human: gene expression; heart; microarray; vascular system; probe;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease; ss.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001MO-US00666.

PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

PT Single exon nucleic acid probes for analyzing gene expression in human  
 PS hearts -  
 CC Claim 4; SEQ ID No 11556; 530bp; English.

CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart. The  
 CC present sequence is one such probe. The probes may be used for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from the human heart via microarrays. By measuring gene expression, the  
 CC probes are useful for predicting, diagnosing, grading, staging,  
 CC monitoring and prognosing diseases of the human heart and vascular system  
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
 CC congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other:

## Alignment Scores:

Pred. No.: 2,91e-71 Length: 567  
 Score: 948.00 Matches: 169  
 Percent Similarity: 97.74% Conservative: 4  
 Best Local Similarity: 95.48% Mismatches: 4  
 Query Match: 38.79% Indels: 0  
 DB: 22 Gaps: 0

US-09-902-772-2 (1-451) x ABA33090 (1-567)

QY 275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294  
 DB 534 CTGACCTCATTTAGGACAGTGGCCAGATCCAGCTTGGGAGTTCTCTCGAGCTCTGCG 475  
 QY 295 AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphelusMetThr 314  
 DB 474 GACAGCTCCAACTCCAGCTCATCAGCTGGGAGGCGCAACGGGAGTTCAAGATGACG 415  
 QY 315 AspProAspGluValAlaIleArgTrpGlyGluArgLysSerLysProAsnMetAsnTyr 334  
 DB 414 GATCCCGACAGAGTGGCCCGCTGGGAGAGCGGAGGCAAAACCAACATGAACTAC 355  
 QY 335 AspLysLeuSerArgAlaIleArgTrpGlyTyrAspLysAsnIleMetThrLysValHis 354  
 DB 354 GATAGCTGACGCGCGCCCTCCCTGACTATGACAAAGAACATCATGACCAAGTTCAT 295  
 QY 355 GlyLysArgTrpAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374  
 DB 294 GGGAGAGCGCTACCGCTCAAGTTGCACTTCCAGGGGATCCCGACGGCTCCAGCCGAC 235  
 QY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394  
 DB 234 CCCCCGAGTCACTCTGTACAGTACCCCTCAGACCTCCCGTACATGGGCTCTATCAC 175  
 QY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414  
 DB 174 GCCACCCACAGAAAGTGAATCTTGTGGGCCCCACCCCTCCAGCCCTCCCGTACATCT 115  
 QY 415 SerSerPhePheAlaAlaProAsnProTyrTTPAsnSerProThrGlyIleTyrPro 434

```

Db      114 TCCAGTTTTTGGTCCCAACCACTGGAATTCACCACTGGGGGTATATACCC 55
QY      435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db      54 AACACTAGGCTCCCAACCACTGGAATTCACCACTGGGGGTATATACCC 4

RESULT 10
AAK14426/c
ID      AAK14426 standard; DNA; 567 BP.
XX
AC      AAK14426;
XX
DT      05-NOV-2001 (first entry)
XX
DE      Human brain expressed single exon probe SEQ ID NO: 14417.
XX
KW      Human; brain expressed exon; gene expression analysis; probe;
KW      microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW      epilepsy; cancer; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157275-A2.
XX
PD      09-AUG-2001.
XX
PE      30-JAN-2001; 2001WO-US00667.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-483446/52.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
PT      brains -
XX
PS      Example 4; SEQ ID NO: 14417; 650bp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      brain. They can be used to measure gene expression in brain cell samples,
CC      which may enable the diagnosis and improved treatment of nervous system
CC      diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC      epilepsy and cancers. The present sequence is one of the probes of the
CC      invention.
XX
SO      Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:
Pred. No.:      2,91e-71      Length:      567
Score:          948.00      Matches:      169
Percent Similarity: 97.74%      Conservative: 4
Best Local Similarity: 95.48%      Mismatches: 4
Query Match:    38.79%      Indels:      0
DB:            22      Gaps:      0

US-09-902-772-2 (1-451) x AAK14426 (1-567)
QY      275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294
Db      534 CTGACCTCATTTAGGACAGGCGGCAAGATCCAGTTGGGAGTCCCTCGGAGCTCTCTGC 475
QY      295 AspSerSerAsnSerAsnCysIleThrTTrpGlnGlyThrAsnGlyIupheLysMetThr 314

```

```

Db      474 GACAGCTCACTCACTCAGCTGCATCACCCTGGGAAGGACCAACGGGGAGTCAAGATGACG 415
QY      315 AspProAspGlyValAlaArgArgTyrGlyGlyUarGlySerLysProAsnMetAsnTyr 334
Db      414 GATCCCGCGAGGAGGGCGCGGCGGTGGGAGAGGGAGCAACCAACCATGANAACATAC 355
QY      335 AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 354
Db      354 GATTAAGCTCAGCCCGCCCTCCGTTACTACTATGACAGAACATCATGACCAGGTCCAT 295
QY      355 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374
Db      294 GGAAGAGGCTACGGCTACAAAGTTGAGACTCCACAGGGAGATCGCCAGGCCCTCAGCCAC 235
QY      375 ProGlnGlySerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394
Db      234 CCCCCGGAGTATCTCTGTACAGTACCCCTCAACACCTCCGTTACATGGGCTCTATCAC 175
QY      395 AlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeuProValThrSer 414
Db      174 GCCCAACCCACAGAGATGAACTTTGTGGCGCCCAACCTCCAGCCCTCCCGTGCATCT 115
QY      415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 434
Db      114 TCCAGTTTTTGGTCCCAACCACTGGAATTCACCACTGGGGGTATATACCC 55
QY      435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db      54 AACACTAGGCTCCCAACCACTGGAATTCACCACTGGGGGTATATACCC 4

RESULT 11
AAK40160/c
ID      AAK40160 standard; DNA; 567 BP.
XX
AC      AAK40160;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human bone marrow expressed single exon probe SEQ ID NO: 14717.
XX
KW      Human; bone marrow expressed exon; gene expression analysis; probe;
KW      microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157276-A2.
XX
PD      09-AUG-2001.
XX
PE      30-JAN-2001; 2001WO-US00668.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488900/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
PT      analyzing gene expression in human bone marrow -
XX
PS      Example 4; SEQ ID NO: 14717; 658bp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human

```

CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukemia and myeloma. The present sequence is one of  
CC the probes of the invention.

XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

#### Alignment Scores:

Pred. No.:	2,91e-71	Length:	567
Score:	948.00	Matches:	169
Percent Similarity:	97.74%	Conservative:	4
Best Local Similarity:	95.48%	Mismatches:	4
Query Match:	38.79%	Indels:	0
DB:	22	Gaps:	0

US-09-902-772-2 (1-451) x AAK40160 (1-567)

```

OY 275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294
DB 534 CTGACCTCATTAAGCGAGCGCCAGATCCAGCTTGGAGTCCCTCCAGCTCCTGCG 475
OY 295 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGlnPheLysMetThr 314
DB 474 GACAGCTCCACTCCAGCTGCATCACCCTGGAGAGCCACCAAGGAGTTCAAGATGACG 415
OY 315 AspProAspGluValAlaAlaArgTrpGlyGlnGlySerLysProAsnMetAsnTyr 334
DB 414 GATCCCGACGAGGAGGCGCGCGCTGGGAGAGCGGAGACCAACCAACATGAACTAC 355
OY 335 AspLysLeuSerArgAlaLeuArgTrpTyrTrpAspLysAsnIleMetThrLysValHis 354
DB 354 GATTAAGCTCAGCGCGCTCCGTTACTACTATGACAAAGACATCATATACCAAGGTCAT 295
OY 355 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374
DB 294 GGAAGAGCGCTACGCTTAAGTTGACTTCACCGGAGATCGCCAGCGCTCCAGCCAC 235
OY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394
DB 234 CCCCCGGAGCATCTCTGTACAGTACCCCTCAGACCTCCGTTACATGGCTCCATAC 175
OY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414
DB 174 GCCACCCACAGAGATGAATTTGTGGGCCCAACCCCTCCAGCCCTCCGTCACATCT 115
OY 415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleTyrPro 434
DB 114 TCCAGTTTGTGTGGCCCCAACCACCACTGGAATTCACCACTGGGGGTATATACCC 55
OY 435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
DB 54 AACACTAGGCTCCGCCACAGCATATGCTTCATCTCGGCGACTTACTAC 4

RESULT 12
AAI20932/c
ID AAI20932 standard; DNA; 567 BP.
AC AAI20932;
XX
XX 12-OCT-2001 (first entry)
DE Probe #10865 for gene expression analysis in human cervical cell sample.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; ss.
XX
XX Homo sapiens.
XX
XX MO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.

```

```

XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25; SEQ ID No 10865; 487P; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SEMP). The present sequence is one such probe. The SEMPs are derived
XX from human HeLa cells. The SEMPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 2,91e-71 Length: 567
XX Score: 948.00 Matches: 169
XX Percent Similarity: 97.74% Conservative: 4
XX Best Local Similarity: 95.48% Mismatches: 4
XX Query Match: 38.79% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-902-772-2 (1-451) x AAI20932 (1-567)
OY 275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294
DB 534 CTGACCTCATTAAGCGAGCGCCAGATCCAGCTTGGAGTCCCTCCAGCTCCTGCG 475
OY 295 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGlnPheLysMetThr 314
DB 474 GACAGCTCCACTCCAGCTGCATCACCCTGGAGAGCCACCAAGGAGTTCAAGATGACG 415
OY 315 AspProAspGluValAlaAlaArgTrpGlyGlnGlySerLysProAsnMetAsnTyr 334
DB 414 GATCCCGACGAGGAGGCGCGCGCTGGGAGAGCGGAGACCAACCAACATGAACTAC 355
OY 335 AspLysLeuSerArgAlaLeuArgTrpTyrTrpAspLysAsnIleMetThrLysValHis 354
DB 354 GATTAAGCTCAGCGCGCTCCGTTACTACTATGACAAAGACATCATATACCAAGGTCAT 295
OY 355 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374
DB 294 GGAAGAGCGCTACGCTTAAGTTGACTTCACCGGAGATCGCCAGCGCTCCAGCCAC 235
OY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394
DB 234 CCCCCGGAGCATCTCTGTACAGTACCCCTCAGACCTCCGTTACATGGCTCCATAC 175
OY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414
DB 174 GCCACCCACAGAGATGAATTTGTGGGCCCAACCCCTCCAGCCCTCCGTCACATCT 115
OY 415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleTyrPro 434

```

Db 114 TCACGTTTTTGTGCTGCCCAACCACTGAAATTCACCACTGGGGGTATATACCCC 55  
 QY 435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451  
 Db 54 AACACTAGGCTCCCGCCAGCATATGCTTCATCTGCGGCACTACTAC 4

RESULT 13  
 AA146176/c  
 ID AA146176 standard; DNA; 567 BP.  
 AC AA146176;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Probe #14862 used to measure gene expression in human placenta sample.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 RM genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200157272-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human placenta -  
 PS Claim 25; SEQ ID NO 14862; 654bp; English.  
 CC The present invention relates to single exon nucleic acid probes (SENP).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.  
 XX  
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,91e-71 Length: 567  
 Score: 948.00 Matches: 169  
 Percent Similarity: 97.74% Conservative: 4  
 Best Local Similarity: 95.48% Mismatch: 4  
 Query Match: 38.79% Indels: 0  
 DB: 22 Gaps: 0

US-09-902-772-2 (1-451) x AA146176 (1-567)

QY 275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294  
 Db 534 CTGACCTCATATAGGCACTGGCGACATCCAGCTTGGCAGTTCCTCGAGCTCCTGCG 475

QY 295 AspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyGlnPheLeuMetThr 314  
 Db 474 GACGCTCCACATCCCTGATCCTGCGGAAGGCAACACGGGAGTTCCAAATGACG 415

QY 315 AspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTyr 334

Db 414 GATCCGACGAGGTGGCCCGCGCTGGGGAAGGGAAGCAACCAACATGATGAC 355  
 QY 335 AsPlysLeuSerArgAlaLeuArgTyrTyrTyrAsPlysAsnIleMetThrLysValHis 354  
 Db 354 GATAGCTCAAGCCCGCCCTCCGTACTACTATGACAAGAACATCATGACCAAGTCCAT 295

QY 355 GlyLysArgGlyAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374  
 Db 294 GGAAGGCGTACGCGCTCAAGATTGACTCCACGGGATGCCAGGCCCTCCACGCCAC 235

QY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394  
 Db 234 CCGCCGAGATATCTCTGTAACAATACCCCTCACAACCTCCGCTACATGGGCTCATAC 175

QY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414  
 Db 174 GCCCACCCACAGAAAGATGAACTTGTGGCGCCCAACCTCCAGGCCCTCCGTCACATCT 115

QY 415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleLeuPro 434  
 Db 114 TCACGTTTTTGTGCTGCCCAACCACTGAAATTCACCACTGGGGGTATATACCCC 55

QY 435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451  
 Db 54 AACACTAGGCTCCCGCCAGCATATGCTTCATCTGCGGCACTACTAC 4

RESULT 14  
 AA106643/c  
 ID AA106643 standard; DNA; 567 BP.  
 XX  
 AC AA106643;  
 XX  
 DT 09-OCT-2001 (first entry)  
 XX  
 DE Probe #6634 used to measure gene expression in human breast sample.  
 XX  
 KW Probe; human; breast disease; breast cancer; development disorder; ss;  
 KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200157270-A2.  
 PD 09-AUG-2001.  
 PF 29-JAN-2001; 2001WO-US00661.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2001-476286/51.  
 XX  
 PT Novel single exon nucleic acid probe used to measuring gene expression  
 XX in a human breast -  
 PS Claim 25; SEQ ID NO 6634; 322bp; English.  
 CC The present invention relates to novel single exon nucleic acid probes.  
 CC The present sequence is one such probe. The probes are useful for  
 CC measuring human gene expression in a human breast sample, where the probe  
 CC hybridises at high stringency to a nucleic acid expressed in the human  
 CC breast. The probes are useful for predicting, diagnosing, grading,  
 CC staging, monitoring and prognosing diseases of the human breast.



Alignment Scores:

Pred. No.:	2,91e-71	Length:	567
Score:	946.00	Matches:	169
Percent Similarity:	97.774%	Conservative:	4
Best Local Similarity:	95.48%	Mismatches:	4
Query Match:	38.79%	Indels:	0
DB:	24	Gaps:	0

US-09-902-772-2 (1-451) x ABS14215 (1-567)

QY	27	LeuAlaAsnProGluSerGlyInnLeuInleuProGlnPheLeuAsnGluLeuLeuSer	294
Db	534	CTGACCTCATTTAGGACAGTGGCCAGATCCAGCTTTGGCATGTTCTCTCTGGACTCTTCG	475
QY	295	AspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyGluPheGlyMetThr	314
Db	474	GACAGCTCCAACTCCAGCTGCATCCCTGGAGAGCCAAACGGGGAGTTCAAGATGACG	415
QY	315	AspProAspGluValAlaArgTrpTrpGlyGluArgGlySerGlyProAsnMetAspTyr	334
Db	414	GATCCCGACGAGGTGGCCCGGGCGTGGGAGAGCGGAGAACCAACCAACATGAATAC	355
QY	335	AspGlyLeuSerAspAlaLeuArgTyrTyrTyrAspGlyAsnIleMetThrGlyValHis	354
Db	354	GATAACTCAGCGCGGCCCTCCGTACTACTATGACAAAGACATCATGACCAAGTCCAT	295
QY	355	GlyGlyArgTyrAlaTyrGlyPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis	374
Db	294	GGGAACGGCTAGCGCTTCAAGATTGCACTTCCACGGGAATGGCCAGGCCCTCCACCCAC	235
QY	375	ProGluGluSerSerMetGlyTyrTyrProSerAspLeuProTyrMetSerSerTyrHis	394
Db	234	CCCCCGAGTCATCTGTGCAAAAGTAAACCCCTGACGCTCCGTTACATGGGGTCTATCAC	175
QY	395	AlaHisProGluGlyMetAspPheValAlaProHisProProAlaLeuProValThrSer	414
Db	174	GCCACACCCACAGAGTGAATCTTGTGGGGCCCCACCTCCAGGCCCTCCCGGTACATCT	115
QY	415	SerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyGlyIleTyrPro	434
Db	114	TCCAGTTTTTTGTGTCGCCCAAAACCATACGTGAATTCACCACTGGGGGATATATACCC	55
QY	435	AspThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451	
Db	54	AATCATGAGCTCCCAACCAATATGCTTCTCATATCGGGCACTTACTAC 4	

Search completed: July 28, 2003, 04:30:21  
Job time : 301.339 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:18:59 ; Search time 60.6835 Seconds  
(without alignments)  
2279.222 Million cell updates/sec

Title: US-09-902-772-2  
Perfect score: 2444  
Sequence: 1 MASTIKALSVSEDSLFE.....IYPNRLPAHMPHSHGTTY 451

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09902772/rmlna.23072003\_093657\_14904/app\_query.fasta.1.1294  
-DB=Issued\_Patents\_NA -OPMT=fastap -SUFFIX=rml -MINMATCH=0.1 -LOOPCL=0  
-LIST=45 -DOCALLIC=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09902772.cgn2\_1.1\_55\_etunal.23072003\_093657\_14904 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2444	100.0	1447	4	US-08-878-177-1
2	2414.5	98.8	1528	4	US-08-878-177-3
3	1624.5	66.5	2938	2	US-08-343-443B-3
4	511	20.9	1752	4	US-09-360-779-1
5	511	20.9	1752	4	US-09-435-335-1
6	449	18.4	1604	1	US-08-306-661B-43
7	449	18.4	1604	5	PCT-US93-06251-9
8	445.5	18.2	2268	3	US-09-344-579-1
9	342.5	14.0	2667	2	US-08-469-412A-1
10	342.5	14.0	2667	4	US-09-021-715-1
11	322	13.2	2064	3	US-08-875-944B-1
12	322	13.2	2064	4	US-09-116-049-3

13	306	12.5	2410	2	US-08-780-835B-1	Sequence 1, Appli
14	306	12.5	2410	4	US-09-303-268-1	Sequence 1, Appli
15	306	12.5	2410	4	US-09-116-049-1	Sequence 1, Appli
16	301.5	12.3	328	2	US-08-343-443B-5	Sequence 5, Appli
17	301	12.3	1905	2	US-09-055-113-2	Sequence 2, Appli
18	291.5	11.9	2266	4	US-09-213-767-1	Sequence 1, Appli
19	273.5	11.2	5510	3	US-09-009-913-3	Sequence 3, Appli
20	261.5	10.7	5427	3	US-09-009-913-2	Sequence 2, Appli
21	261.5	10.7	5667	3	US-09-009-913-4	Sequence 4, Appli
22	261.5	10.7	5667	2	US-08-469-412A-6	Sequence 6, Appli
23	234.5	9.6	2544	4	US-09-021-715-6	Sequence 6, Appli
24	231.5	9.5	848	3	US-09-009-913-338	Sequence 338, App
25	231.5	9.5	2280	3	US-09-009-913-8	Sequence 8, Appli
26	231.5	9.5	2498	3	US-09-009-913-10	Sequence 6, Appli
27	231.5	9.5	2498	4	US-09-009-913-10	Sequence 10, Appli
28	229	9.4	852	4	US-09-020-956-44	Sequence 44, Appli
29	229	9.4	852	4	US-09-030-607-44	Sequence 44, Appli
30	229	9.4	852	4	US-09-605-785-44	Sequence 44, Appli
31	229	9.4	852	4	US-09-439-313-44	Sequence 44, Appli
32	229	9.4	852	4	US-09-352-616A-44	Sequence 44, Appli
33	229	9.4	852	4	US-08-232-149A-44	Sequence 44, Appli
34	227.5	9.3	1920	1	US-08-746-789A-1	Sequence 1, Appli
35	211.5	8.7	2975	1	US-08-368-281-1	Sequence 1, Appli
36	211.5	8.7	3240	1	US-08-368-281-3	Sequence 3, Appli
37	181	7.4	65042	4	US-09-784-316-3	Sequence 50, Appli
38	166	6.8	1364	5	PCT-US93-06251-65	Sequence 65, Appli
39	166	6.8	50937	4	US-09-428-517-1	Sequence 1, Appli
40	118.5	4.8	5194	1	US-09-599-652-1	Sequence 1, Appli
41	117	4.8	5194	2	US-08-642-846-1	Sequence 1, Appli
42	117	4.8	5194	4	US-09-264-604-1	Sequence 1, Appli
43	115.5	4.7	2626	1	US-08-156-020-5	Sequence 5, Appli
44	115.5	4.7	2626	1	US-08-156-020-3	Sequence 3, Appli
45	114.5	4.7	2626	1	US-08-156-020-3	Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-08-878-177-1  
Sequence 1, Application US/08878177  
Patent No. 6294354  
GENERAL INFORMATION:  
APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et  
TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
FILE REFERENCE: Chugai Selyaku Kabushiki Kaisha 5001  
CURRENT FILING DATE: 1997-06-18  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 1  
LENGTH: 1447  
TYPE: DNA  
ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA  
US-08-878-177-1

## Alignment Scores:

Pred. No.: 4.05e-247 Length: 1447  
Score: 2444.00 Matches: 451  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-902-772-2 (1-451) x US-08-878-177-1 (1-1447)

Qy 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
Db 63 ATGCAAGCACCTATTAAAGAGACATTATCACTGTGAGTGAAGACCACTCTTTTGGAG 122  
Qy 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
Db 123 TGTGCTTACGGATGCGCCACCTTGCAGAAAGACAAATGACAGCCTCTCTTCACGAGAA 182

QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60  
 DB 183 TATGGCAACATCAAGATGAGAGCCGCGGTTCCCGACAGAGCTGTTATCATCACAGCC 242  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 DB 243 CCGGCCAGAGTTACCATTAAGATGAGTGAACCCAAACAGGTTAATGGGTCAAGGAAT 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 DB 303 TCACCTGATGACATGCAGCGTGGCAAAAGAGGAAATGGTTAGCACTTCAGCAATGTT 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120  
 DB 363 GGGATGAACATATGGAAGCTCATGGAAGAGACATATTCGCGCTCCAAATATGACAAAC 422  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140  
 DB 423 AATGACGAAGAGATTATGTGCCAGACATCTTACCTTATGAGACACAGACCATGTACGG 482  
 QY 141 GlnTyrLeuGlnTyrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 DB 483 CAGTGGCTGAGTGGGAGTGAAGAGATGATGCTTCCAGACCGTGCATCTTGTTCTTC 542  
 QY 161 GluAsnIleAspGluLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 DB 543 CAGAACATTCATGGGAAAGAGTGTGTAATGACAAAGATGACTTCCAGACACTCAGC 602  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgValAla 200  
 DB 603 CCGAGCTATACCAAGATATCTCTGCTGACACTTACACTTACCTCAGAGAGAGAGAGGCC 662  
 QY 201 ThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThrArg 220  
 DB 663 ACTTTATTTTCCAAATACATCAGTTTACCCAGAAAGCAAGCAAGAAATACAAACAGG 722  
 QY 221 ProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThrSerHisSerHisProThr 240  
 DB 723 CCGAGTTACCTTATAGCAAGAGAGAGATCAGCTGCGAGTGCAGACCCCTCCACT 782  
 QY 241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgPro 260  
 DB 783 CAGTCAAAAGCTACCAACATCATCTTCAACAGTCCCAAAACAGAACGCCAGCTCT 842  
 QY 261 GlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280  
 DB 843 CAGTTAGATCCTTATCAGATTCTTGGACGACAGCCGCTTCCAAATCCAGGAGT 902  
 QY 281 GlyGlnIleGlnLeuThrPheGlnPheLeuLeuLeuLeuLeuSerAspSerSerAsn 300  
 DB 903 GGGCAGATACAGCTATGAGAGTTCACGTGAGACCTTCTCGGACAGCTCCAACTCCAAC 962  
 QY 301 CysIleThrTrpGluGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAla 320  
 DB 963 TGCATACCTCGGAGGACCAAAATGGGAGTTCAAATGACAGACCTCGATGAGTGGCT 1022  
 QY 321 ArgArgTyrGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAla 340  
 DB 1023 CCGCGTTGGGAGAGAGAAAGCAAACTAACATGACATGACAACTCAGCCGTGCA 1082  
 QY 341 LeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360  
 DB 1083 CTTCCGTACTACTATGACAAAAATATATGACTAAAGTTCATGCTAAAGCTATGCTCTAC 1142  
 QY 361 LysPheAspPheHisGlyLysIleAlaGlnAlaLeuGlnProHisProProGluSerSerMet 380  
 DB 1143 AATTTGATTTCCACGGAATCGCTCAGGCCCTCCACCCCTCACCAATATCATCATG 1202  
 QY 381 TyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMet 400  
 DB 1203 TACAATATACCATGAGACCTCCCTACATGAGTTCCTACATGACACCCCAAGAAAGT 1262

QY 401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAlaAla 420  
 DB 1263 AACTTTTACTCTCCCATCCCTCCCTTTCGCGGTACCTCATCCACTTTTTCGTGCC 1322  
 QY 421 ProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrProAsnThrArgLeuProAla 440  
 DB 1323 CCTAATCCATATGGAATTCACCAACATGAGGAGATTCACCAATACCAAGCTGCAGCT 1382  
 QY 441 AlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 DB 1383 GCTCATATGCTTCCCATCTTGGCAGCTACTAC 1415  
 RESULT 2  
 US-08-878-177-3  
 ; Sequence 3, Application US/08878177  
 ; Patent No. 6294354  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwanoto et  
 ; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
 ; FILE REFERENCE: Chugai selyaku kabushiki kaisha 5001  
 ; CURRENT APPLICATION NUMBER: US/08/878,177  
 ; CURRENT FILING DATE: 1997-06-18  
 ; NUMBER OF SEQ. ID NOS.: 7  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ. ID NO. 3  
 ; LENGTH: 1528  
 ; TYPE: DNA  
 ; ORGANISM: c-ery gene, chicken DNA  
 US-08-878-177-3  
 Alignment Scores:  
 Pred. No.: 5,57e-244 Length: 1528  
 Score: 2414.50 Matches: 450  
 Percent Similarity: 94.14% Conservative: 0  
 Best Local Similarity: 94.14% Mismatches: 1  
 Query Match: 98.79% Indels: 27  
 DB: 4 Gaps: 1  
 US-09-902-772-2 (1-451) x US-08-878-177-3 (1-1528)  
 QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 DB 63 ATGGCAAGCACTATTAAGAGAGATGATGAGTGAAGACAGTCTCTGTTGAG 122  
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 DB 123 TGTGGCTAGGAGTGGCCGACCTTGAAGACAGAAATGACAGCTCTTCCATGAA 182  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60  
 DB 183 TATGGCAACATCAAGATGAGAGCCGCGGTTCCCGACAGAGCTGTTATCATCACAGCC 242  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 DB 243 CCGGCCAGAGTTACCATTAAGATGAGTGAACCCAAACAGGTTAATGGGTCAAGGAAT 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 DB 303 TCACCTGATGACATGCAGCGTGGCAAAAGAGGAAATGGTTAGCACTTCAGCAATGTT 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120  
 DB 363 GGGATGAACATATGGAAGCTCATGGAAGAGACATATTCGCGCTCCAAATATGACAAAC 422  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140  
 DB 423 AATGACGAAGAGATTATGTGCCAGACATCTTACCTTATGAGACACAGACCATGTACGG 482  
 QY 141 GlnTyrLeuGlnTyrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 DB 483 CAGTGGCTGAGTGGGAGTGAAGAGATGATGCTTCCAGACCGTGCATCTTGTTCTTC 542

QY 161 GlnAsnIleAspGlyLeuGluCysLeuMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 543 CAGAACTGTGATGGAGAAAGTTGTGTAATGACCAAGATGCTCCAGAGCTCAGC 602  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg 198  
 Db 603 CCGAGCTATACGAGATATCTCTCTACACCTACATACCTCAGAGAGACTCTCTT 662  
 QY 198 ----- 198  
 Db 663 CCACATTGACTCAGATGATGTTGATAGAGCCCTTACAAAACCTCCACGGTTATGAT 722  
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213  
 Db 723 GCTAGAAACACAGAGAGAGCCACTTTATTTTCCAAATACATCAGTTTACCCGAAAGCA 782  
 QY 214 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTyr 233  
 Db 783 ACGCAAGAAATTAACACAGAGCCAGATTTTACCTTATGAGCAAGGAGAGATCAGCGTGG 842  
 QY 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 253  
 Db 843 ACGAGTACAGCCATCCCTCCTCAGTCAAAAGCTACCAACCATCATCTTCAACAGTCCC 902  
 QY 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273  
 Db 903 AAAACAGAACAGCCGCTCTCTCAGTTAGATCCTTATCAGATCTTGGACCGACGACGAC 962  
 QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGluLeuLeu 293  
 Db 963 CGTCTTCAAAATCCAGGAGAGTGGGACATACAGCTATGAGCAGTCTCTAGAGACTCTCTG 1022  
 QY 294 SerAspSerSerAsnSerAsnGlyIleThrTyrGluGlnArgGlySerLysProAsnMet 313  
 Db 1023 TCGGACAGCTCCAACTCACTGATCCTCGGAGGAGGACAAATGGGAGTTCAAGATG 1082  
 QY 314 ThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysProAsnMetAsn 333  
 Db 1083 ACAGACCTCATGAAAGGCTCGGCGTTGGGGAGAGAGGAAAGCAAACTTACATTAAC 1142  
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353  
 Db 1143 TATGACAACTCAGCCCTGACCTCGCTACTACTATGACAAATAATATTATGACTAAGTT 1202  
 QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisIleAlaGlnAlaLeuGlnPro 373  
 Db 1203 CATGGTAACGCTATGCTCAAAATTTGATTTCACGGAATCGCTCAGGCCCTCCACCT 1262  
 QY 374 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393  
 Db 1263 CACCCCTCCAGATCATCATCATGATCAATACCCATCAGACCTCCCTTACATGAGTTCTTAC 1322  
 QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeuProValThr 413  
 Db 1323 CATGCAACCCCAAGAGATGAATTTGTAGCTCCCATCCCTCGCTTGGCCGTAACC 1382  
 QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyGlyIleTyr 433  
 Db 1383 TCATCCAGCTTTTGTGCTGCCCTTAATCATCATGGAATTCACCACTGGAGGATCTTAC 1442  
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 Db 1443 CCCAATACAGGCTGCCAGCTCATATGCTTCCATCTTGGCACTTACTTAC 1496  
 RESULT 3  
 US-08-343-443B-3  
 : Sequence 3, Application US/08343443B  
 : Patent No. 5968734  
 : GENERAL INFORMATION:  
 : APPLICANT: Aurias, Alain  
 : APPLICANT: Delattre, Olivier  
 : APPLICANT: Desmaze, Chantal  
 : APPLICANT: Melot, Thomas

: APPLICANT: Peter, Martine  
 : APPLICANT: Ploongastel, Beatrice  
 : APPLICANT: Thomas, Gilles  
 : APPLICANT: Zucman, Jessica  
 : TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF  
 : TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL  
 : TITLE OF INVENTION: TRANSLOCATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS  
 : TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID  
 : TITLE OF INVENTION: TRANSLOCATIONS  
 : NUMBER OF SEQUENCES: 129  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Weisner & Associates  
 : STREET: 230 South Fifteenth Street  
 : City: Philadelphia  
 : STATE: PA  
 : COUNTRY: USA  
 : ZIP: 19102  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: AEDIT 1.0 DOS text editor  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/343,443B  
 : FILING DATE: 18-NOV-1994  
 : CLASSIFICATION: 514  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/FR93/00494  
 : FILING DATE: 19-MAY-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: FR 92/06123  
 : FILING DATE: 20-MAY-1992  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Weisner, Gerard J.  
 : REGISTRATION NUMBER: 19,763  
 : REFERENCE/DOCKET NUMBER: 989,6121P  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 215-875-8383  
 : TELEFAX: 215-875-8394  
 : INFORMATION FOR SEQ ID NO: 3:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2938 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 143..1498  
 : US-08-343-443B-3  
 : Alignment Scores:  
 : Pred. No.: 1,696-160 Length: 2938  
 : Score: 1624.50 Matches: 306  
 : Percent Similarity: 79.43% Conservative: 57  
 : Best Local Similarity: 66.96% Mismatches: 83  
 : Query Match: 66.47% Indels: 11  
 : Gaps: 7  
 : US-09-902-772-2 (1-451) x US-08-343-443B-3 (1-2938)  
 QY 1 MetalaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 143 ATGGACGGGACTATTAAAGAGGCTCTGTGGTGGAGCAAGCAAGTCCCTCTTGAC 202  
 QY 21 CysAlaTyrGly--SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39  
 Db 203 TCAGCGTAGCGAGGAGCGAGCCATCTCCCAAGGCCAGCAAGTGCCTGGGAGATCCT 262  
 QY 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGln 59  
 Db 263 GACTACGGGAGGCCCAAGAGATCAACCCCTCCACACAGCAGGAGTGGATCAATCAG 322  
 QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79

Db	323	---	CCAGTAGGGGTCAACGTCAAGCGGAGTAT-----GACCATAAGTATGATCCAGG	373
QY	80	AsnSerProAspAspCysSerValAlaIalysGlyLysMetValSerSerAspAsn	99	
Db	374	GAGTCTCCGTGGAGCTGACGCGTTACCAATGCAGCAACCTGGTGGCGGAGCGAGTCC	433	
QY	100	ValGlyMetAsnTyrGlySerTyrMetGlnGlyLysHis--LLeProProAspMet	118	
Db	434	AACCCCATTAATACACAGCTATATNGAGAGAAATAATGGCCCCCTCTCCCAACATG	493	
QY	119	ThrThrAsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHis	138	
Db	494	ACCACCACAGGAAGAGAGTACTCGCCCCGACGCCACACACTGTGGACAGGAGCAT	553	
QY	139	ValArgGlnTrpLeuGlnTrpPheValValIlysGlyTyrGlyLeuProAspValAspIleLeu	158	
Db	554	GTCAGGCAATGGCTGAGTGGGCCATAAAGAGTATAGCTTATGGAGATCGACATCC	613	
QY	159	LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg	178	
Db	614	TTTTTCCAGAACTGATGGACGAAGAACTGTGTAATGAACAAAGAGAGACTTCCTCCGC	673	
QY	179	LeuThrProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgValArg	198	
Db	674	GCCACACACCTTACACAGGAGGCTGTGTCACTCACTCACTTACCTCAGGAGAA---	730	
QY	199	GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThr	218	
Db	731	---AGTTCACTGCTGGCCTATATATCAACCTCCACACGACCAATCCTCAGATTGAGT	787	
QY	219	ThrArgProAspLeuProTyrGlnAlaIalArgArgSerAlaThrPheSerHisSerHis	238	
Db	788	GTCAAAGAGAACCTCTTATGACTACGACAGAAAGGAGCAAGGGGCAATAACATGAAT	847	
QY	239	ProThrGlnSerLeuAlaThrGln---ProSerSerSerThrValProLysThrGluAsp	257	
Db	848	TCTGGCCTCAACAAAAGTCTCCCTGGAGGGGCAACAAACATCACTAAGATACAGAG	907	
QY	258	GlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn	277	
Db	908	CACACGCCCCACAGATCCGTATCGATCCGTGGCCCCACACGACGATGCCCTACGCAAC	967	
QY	278	ProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGlnLeuLeuSerAspSerSer	297	
Db	968	CTTGAAGGGGACAGTCCAGTGTGGCAATTCCTCTGAGTGTCTCCGACACGGCC	1020	
QY	298	AsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGlyPheLysMetThrAspProAsp	317	
Db	1028	AACGCGACGTGATCACTCGGAGGGGACCAAGGGGAGTTCAAATATGACGAGCCCGAT	1088	
QY	318	GluValAlaIalArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTyrAspLysLeu	337	
Db	1088	GAGGTGGCCAGGCGCTGGGGCCAGCGAAACCAACCCAACTGTGAATTAACGCAAGCTG	1144	
QY	338	SerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArg	357	
Db	1148	AGCGGGCGCTCCGTTATTACTATGATAAAACATTATGACCAAACTGCACGGCAAAAG	1200	
QY	358	TyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGln	377	
Db	1208	TATGCTTCAAAATTTGACTTCACAGCATTCGCCAGGCTCTGCAGCCACATCCGACGAG	1266	
QY	378	SerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisPro	397	
Db	1268	TGCTCATAGTAAAGTACCTTCCTGACATCTCTTACATCTCTTCACATGCCCCACCG	1322	
QY	398	GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhe	417	
Db	1328	CAGAAAGTGAAGTCTTGCTCCCTCCCATCATCTCTCATGCGCTGTACTCTCTCCACTTC	1388	
QY	418	PheAlaAlaProAsnProTyrThrAsnSerProThrGlyGlyTyrLeTyrProAsn-----	435	

```

Db      1388  TTGGAGCGGCATCATCAATACTGAGACCTCCGCCAGGGGGGGAATCTACCCAAACCCCAAC 1447
QY      436  ---ThraGleuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db      1448  GTCCCCCGCATCTTAACACCCACGCTGCTTACACTTAAGGAGCTACTAC 1498

RESULT 4
US-09-360-779-1
: Sequence 1, Application US/09360779
: Patient No. 6268216
: GENERAL INFORMATION:
: APPLICANT: Denetis, Evan S. 46
: APPLICANT: Fyodorov, Dmitry V.
: APPLICANT: Hendicks, Timothy J.
: TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
: TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases
: FILE REFERENCE: CASE-030828
: CURRENT APPLICATION NUMBER: US/09/360,779
: CURRENT FILING DATE: 1999-07-26
: EARLIER APPLICATION NUMBER: 60/094,264
: EARLIER FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1752
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (112)..(1131)
US-09-360-779-1

Alignment Scores:
Pred. No.: 7.4e-44 Length: 1752
Score: 511.00 Matches: 123
Best Local Similarity: 50.69% Conservative: 23
Query Match: 42.71% Mismatches: 63
DB: 20.91% Indels: 79
Gaps: 10

US-09-902-772-2 (1-451) x US-09-360-779-1 (1-1752)
QY      209  ValTYrPRoGluAlaThrGlnAlaArgLeuThrArgProAspLeuProTYrGluGlnAla 228
Db      369  GTGCCCGCCCGTGGCGGTCCCGCCCATCTCCACCGCCCA----- 407
QY      229  ArgArgSerAlaTrpThrSerHisSerHisSerProThrGlnSerLysAlaThrGlnProSer 248
Db      408  -----GTCCCGCGAGCGATGAGACAGAGGCGGACCTCCAGCCCTGCT 452
QY      249  SerSerThrValProLysThrGlnAspGlnArg-----Pro 260
Db      453  GATCAACATGTACTACACAGATCCCGTGGAGATGCTTTTAAAGAGAGGAAGAGCC 512
QY      261  GlnLeuAspProTYrGlnLeuLeuGlyProThrSerSerArgLeuAlaAsnProGlySe 280
Db      513  GAGCTGGGGGGCG-----CTGAGCGCCGCGGTACAGAAA-----GCGAG 551
QY      280  rGlyGlnLeuGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSerAs 300
Db      552  CGGGCAGATCCAGATTGTGGAGTTTCTACTGTGAGCTGCTGCAGACCGCGGAACGCGCG 611
QY      300  nCysIleThrTrpGluGlyThrAsnGlyGlnPheLysMetThrAspProAspGluValAl 320
Db      612  CTGCATCCGCTGGGGGGGGCGCCACGGCGAATTCAAGCTCACCGACCCCGACGAGGTGGC 671
QY      320  aArgArgTrpGlyLysArgLysSerLysProAsnMetAsnTYrAspLysLeuSerArgAl 340
Db      672  GCGACGCTGGGGCGCGCCAGAGACAGACCCCAATATGAATAACGCAAGCAAGTAACTGAGC 731
QY      340  aLeuArgTYrTYrAspLysAsnIleMetThrLysValHisGlyLysArgTYrAlaTY 360
Db      732  ACTCGCGTACTACAGCAAAACATCATATAGCAAGGTGACGCGCAAGCGCTACGCTA 791

```

QY 360 rlyspheasphhisgllylealaglalaLeuGlnProHisPro----- 375  
 Db 792 CCGCTTGACTTCCAGGAGCGGACAGGCTTGCCACACCCGCGCAGCCGCGC 851  
 QY 376 -----ProGlnSerSerMetTyrlYsTyPr 384  
 Db 852 CGCGCGTGGCGCGCGCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 911  
 QY 384 oSerAspLeu-----ProTyMetSerSerTyrlHisAlaHisProGlnLy 399  
 Db 912 GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950  
 QY 399 smetAspHeValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAl 419  
 Db 951 ACTGACCTTATGACGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 999  
 QY 419 aAlaProAsnProTyrlTyPrAsnSerPro-----ThrGlyG1 431  
 Db 1000 -----TACTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1046  
 QY 431 ylleTyPrProAsnThrArgleuPro-----AlaAlaHisMetPr 444  
 Db 1047 GCTCTACCCCAACCCCGGCGCTTGCGAGCCCGCTGCGCGCGCGCGCGCGCG 1106  
 QY 444 oSerHisLeuGlyThrTyPr 451  
 Db 1107 TTCGCACTTGGGGGCTCATTTAT 1128

## RESULT 5

US-09-435-335-1  
 : Sequence 1, Application US/09435335  
 : Patent No. 6384204  
 : GENERAL INFORMATION:  
 : APPLICANT: Demeris, Evan S.  
 : APPLICANT: Eyodoro, Dmitry V.  
 : APPLICANT: Hendricks, Timothy J.  
 : TITLE OF INVENTION: Reagents and Methods for the screening of compounds  
 : FILE REFERENCE: CASE-04027  
 : CURRENT APPLICATION NUMBER: US/09/435,335  
 : EARLIER FILING DATE: 1999-11-05  
 : EARLIER APPLICATION NUMBER: 09/360,779  
 : SOFTWARE: Patent In Ver. 2.0  
 : SEQ ID NO 1  
 : LENGTH: 1752  
 : TYPE: DNA  
 : ORGANISM: Rattus norvegicus  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: (112)..(1131)  
 : US-09-435-335-1

## Alignment Scores:

Pred. No.: 7.4e-44 Length: 1752  
 Score: 511.00 Matches: 123  
 Percent Similarity: 50.69% Conserved: 23  
 Best Local Similarity: 42.71% Mismatches: 63  
 Query Match: 20.91% Indels: 79  
 Gaps: 10

US-09-902-772-2 (1-451) x US-09-435-335-1 (1-1752)  
 QY 209 ValTyPrProGlnAlaThrGlnArgyleThrThrArgProAsnLeuProTyrlGlnGlnAla 228  
 Db 369 GTGGCG 407  
 QY 229 ArgArgSerAlaTrpTrpSerHisSerHisProHisGlnSerLysAlaThrGlnProSer 248  
 Db 408 -----GTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 452

QY 249 SerSerThrValProLysThrGluAspGlnArg-----Pro 260  
 Db 453 GATCAACATGTACTTACCCAGATCCCGTGGAGATGCTTTTAAAGAGGAGAGAGCC 512  
 QY 261 GlnLeu-AspProTyrlGlnLeuGlyProThrSerSerArgleuAlaAsnProGlySe 280  
 Db 513 GAGCTGGGGGCGG-----CTGAGCGCCGCGGCTACAGAA-----GGCAG 551  
 QY 280 rGlyGlnLeuGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSerAs 300  
 Db 552 CGGCGAGATCCAGTTGTTGGCGAGTTTCTACTGAGCTGTCGACACCGCGCGCGCGG 611  
 QY 300 nGySILeThrTrpGlnGlyThrAsnGlyLuprGlyMetThrAspProAspGluValAl 320  
 Db 612 CTGCATGCGCGTGGGAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 671  
 QY 320 AARGATrPGLyLuarGlySerLysProAsnMetAsnTyrlAspLysLeuSerArgAl 340  
 Db 672 GCGAGCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 731  
 QY 340 AleuArgTyrlTyPrAspLysAsnLleMetThrLysValHisGlyLysArgTyrlaTy 360  
 Db 732 ACTGCGCTACTACTACGACAAACATCATGACGAGGTCGACGCGCAAGCGCTACGCTA 791  
 QY 360 rlyspheasphhisgllylealaglalaLeuGlnProHisPro----- 375  
 Db 792 CCGCTTGACTTCCAGGAGCGGACAGGCTTGCCACACCCGCGCAGCCGCGCAGCC 851  
 QY 376 -----ProGlnSerSerMetTyrlYsTyPr 384  
 Db 852 CGCGCGTGGCG 911  
 QY 384 oSerAspLeu-----ProTyMetSerSerTyrlHisAlaHisProGlnLy 399  
 Db 912 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950  
 QY 399 smetAspHeValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAl 419  
 Db 951 ACTGACCTTATGACGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 999  
 QY 419 aAlaProAsnProTyrlTyPrAsnSerPro-----ThrGlyG1 431  
 Db 1000 -----TACTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1046  
 QY 431 ylleTyPrProAsnThrArgleuPro-----AlaAlaHisMetPr 444  
 Db 1047 GCTCTACCCCAACCCCGGCGCTTGCGAGCCCGCTGCGCGCGCGCGCGCGCG 1106  
 QY 444 oSerHisLeuGlyThrTyPr 451  
 Db 1107 TTCGCACTTGGGGGCTCATTTAT 1128

RESULT 6  
 US-08-306-691B-43  
 : Sequence 43, Application US/08306691B  
 : Patent No. 5734039  
 : GENERAL INFORMATION:  
 : APPLICANT: Calabretta, Bruno  
 : APPLICANT: Skorski, Tomasz  
 : TITLE OF INVENTION: ANTISENSE  
 : TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
 : NUMBER OF SEQUENCES: 55  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Seidel, Gonda, Lavorina & Monaco, P.C.  
 : STREET: Two Penn Center, Suite 1800  
 : CITY: Philadelphia  
 : STATE: Pennsylvania  
 : COUNTRY: U.S.A.  
 : ZIP: 19102  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette, 3.50 inch, 720 KB  
 : COMPUTER: IBM PS/2  
 : OPERATING SYSTEM: MS-DOS

```

SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1604 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-306-691B-43

Alignment Scores:
Pred. No.: 2,1e-37 Length: 1604
Score: 449.00 Matches: 114
Percent Similarity: 43.88% Conservative: 51
Best Local Similarity: 30.32% Mismatches: 89
Query Match: 18.37% Indels: 122
Gaps: 9

US-09-902-772-2 (1-451) x US-08-306-691B-43 (1-1604)
QY 119 ThrAsnGlnArgValIleValProAlaAspProThrLeuTyrSerThrAspHis 138
Db 447 ACTAAGAGACAGACACTGGGGGATCCCAAGAGCCCGGACGTGAGACAAACCAT 506
QY 139 ValArgGlnTrpLeuGluTrpAlaValIleGluTyrGlyLeuProAspValAspIleLeu 158
Db 507 GTTCGGGAGCTGGGTGATGGGCTGTGAATGAAATTCAGCCGTGAAGGTGAGACTTCAG 566
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
Db 567 AAGTTTC---TGATGATGATGAGAGAGCCCTCTGGCCGTGGTAAGACTCTTCTTCAG 623
QY 179 LeuThrProSerTyrAsnAlaAspIleLeuSerHisIleuHisTyrIleuArgIleu 197
Db 624 CTGGCCCAAGACTTGTGGGACATCTTATGGACATCTAGAGATCTGACAGAAAGAG 683
QY 198 ---ArgGlyAlaThrPheIlePheProAsnThr 207
Db 684 GATGTGAACCATATCAAGTTAATGAGTCAACCCAGCCATATCAGAAATCCCGTATACC 743
QY 207 --- 207
Db 744 TCGGATTACTTCACTTACCTATGATATGAGCATGCCAGTGTTCACCATCGAGATTC 803
QY 208 ---SerValIleProGluAlaThrGlnArg 216
Db 804 TCAGAGCCCAAGCTTCATCAGAGAGTCTTATGACAGCTCATCCATCCATCAGTCGGAAGAG 863
QY 217 Ile--- 217
Db 864 CTCTCTCCCTCAAGATAGATGAGATACCCCTGGTCAATTCGAGAGACCCCTCTCAG 923
QY 217 --- 217
Db 924 ACAGACACCTTGAGATGACTACTTGTATCAACAAGAGTGTGACCCAGACAGACAC 983
QY 218 ---ThrThrArgProAspLeuProTyrGluGlnAlaArgAspSer--- 231
Db 984 ATGTGATGGGAGAGACCAAGTCGTGTAACCTCGGGGGCCAGAGACTCTTTTGAAGACATA 1043

```

```

QY 232 ---AlaThrPheSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSer 250
Db 1044 GAGAGCTACGATGATGTGATGCGCTCACCAGAGCTCGAGAGCCAGCATCATCTTTCAC 1103
QY 251 ThrValProLys---ThrLysAsp--- 257
Db 1104 AGCCTGAGGAGTGTCCCTCCTATGACAGCTTCGACATCAGAGACATATCCGCTGCCCTG 1163
QY 258 ---GlnArgProGlnLeuAsp 263
Db 1164 CCACACCAAGCCCAAGGACCTTCACAGCATATGTGGGGAGCCGTGACATCAT 1223
QY 264 ProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIle 283
Db 1224 AAGGACAGAGCTGTCAATCTGCTGCTGCCCTAGCTGCTGATACACAGCACTGACCAATC 1283
QY 284 GlnLeuTrpGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsnCysIleThr 303
Db 1284 CAGCTATGAGCAGTTTCTTCTGGAATTAATCACTGATTAATCTCTGACGTCTTTATCAGC 1343
QY 304 TyrGlnGlyThrAsnGlnGlyIlePheLysMetThrAspProAspGluValAlaArgArgTrp 323
Db 1344 TGGACAGAGATGGCTGGGATTCAACTTCTGACCCAGATGAGGTGGCCAGAGATGG 1403
QY 324 GlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyr 343
Db 1404 GGAAGAGGAGAAACAAACCTAGATGATTAATGAGAAATGAGCGCTGCGCTACGCTAC 1463
QY 344 TyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyrLysPhe--- 362
Db 1464 TATTACGACAAACAAATCATCATCCACAGACGCGGGAAGACGTAACGTGTCTTTGTG 1523
QY 363 ---AspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu 377
Db 1524 TGTGACTG---CAGAGCTCTGCGGGTACACCCCTGAG 1559

RESULT 7
PCT-US93-06251-9
Sequence 9, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1604 base pairs
TYPE: nucleic acid

```

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US93-06251-9

## Alignment Scores:

Pred. No.:	2,1e-37	Length:	1604
Score:	449.00	Matches:	114
Percent Similarity:	43.88%	Conservative:	51
Best Local Similarity:	30.32%	Mismatches:	89
Query Match:	18.37%	Indels:	122
	5	Gaps:	9

US-09-902-772-2 (1-451) x PCT-US93-06251-9 (1-1604)

```

QY 119 ThThrsnsluargvalillevalprolaasprothleutrpserthrasphis 138
    ||| :|||:|||||:||||| ||||| |||||: |||
Db 447 ACTAAAGACAGACAGACAGGAGATCCAAAGACCCCGGAGAGGACAGAAACCCAT 506
    ||||| |||||: ||||| |||||: |||

QY 139 ValArglnlrpleuuglutrpalavallysglulrleuproaspvalaspileu 158
    ||||| |||||: ||||| |||||: |||
Db 507 GTTCGGGAGCTGGGTGATGGGCTGTGAATGAAATTCAGCTGAAGGTGAGACTTCCAG 566
    ||||| |||||: ||||| |||||: |||

QY 159 Leupheglasnlleaspllylsglulrleucyslysmethrlyaspaasphelinar 178
    ||| :|||:|||||:||||| ||||| |||||: |||
Db 567 AAGTTC---TGATGAATGAGACAGAGCCCTCTGGCCCTGGGTAAAGACCTCTTCGAG 623
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 179 LeuthrproserlyrasnalaaspileuleuSerHisleuHstlyrleuArglu--- 197
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 624 CTGGCCCAACACTTGTGGGACATCTTATGGAAACATTAAGATCCCTGCAGAAAGAG 683
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 198 -----ArgglyAlaThrPheleuPheproasnThr----- 207
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 684 GATGTGAACCATATCAATGTAATGAGTCAACCCAGCTATTCAGAAATCCCGTATACC 743
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 207 ----- 207
Db 744 TCGGATTACTTCAATAGCTATGATGATAGCATGCCAGTGTGTCCACATCGAGATTG 803
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 208 -----SerVallyrProgluAlaThrArg 216
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 804 TCAGAGCCAGCTTCATCAGACAGTCTATCAGACGCTCATCCATCCAGCTGGAGAG 863
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 217 Ile----- 217
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 864 CTCTCTCCCTCAAGTATGAGAAATGACTACCCCTCGTCTATTCGAGACCTCTCCAG 923
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 217 ----- 217
Db 924 ACAGACACCTTGAGAAATGACTACTTTCATCAACAAAGACGTCACCCAGACAC 983
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 218 -----ThrThrArgProaspLeuProTyrGlulAlaIleArgSer--- 231
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 984 ATGTGCATGGGAGGAGACCACTCGTGAACCTCGGGGCGCAGACACTCTTTGAAGACATA 1043
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 232 ---AlatrpHrSerHisSerHisProThrGlnSerlysalathrGlnProSerSer 250
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 1044 GAAAGCTACGATAGTTGTGATCGCCCTCACCAGTCCGAGAGCCAGTCATCTTTCAC 1103
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 251 ThrValProlys-----ThrGluasp----- 257
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 1104 AGCTGAGCGGTTCCTCCCTCTATGACAGTTCGAGTCAGAGACATATCGCGGCGCTG 1163
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 258 -----GlnArgProGlnLeuasp 263
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 1164 CCCAACCAAGCCCAAGGACACCTTCAGAGCATATGTGGGAGCGGTGCTGACCAAT 1223
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 264 ProTyrGlnIleleuGlyProThrSerSerArgleuAlaIleProGlySerGlyGlnIle 283
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 1224 AAGGACAAAGCTTCATCTCTGCTGCTGCGCTAGCTGAGTACACAGGACATGGAACCAATC 1283
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 284 GlnleuTrpGlnPheleuLeuGluLeuLeuSerAspSerSerAsnSerAsnGlyIleThr 303
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

```

```

Db 1284 CAGCTATGCGAGTTTCTTGTGAATTAATCTGACTGATTAATCTGCTTTTATCAGC 1343
QY 304 TrpGlnGlyThrAsnGlyupheliyMetThrAspProaspGluValAlaArgTyr 323
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 1344 TGACAGAGAGATGGCTGGGAATTCAACTTCTGACCCAGATGAGTGGCCAGGATGG 1403
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 324 GlyGluArglySerlyProAsnMetAsnTyrAspIlyleuSerAlaIleArgTyr 343
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 1404 GGAAGAGGAAAAAACAACCTAGATCAATTAATGAGAACTGAGCCGTGAGCTTAC 1463
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 344 TyrTyrAspIlysnlleMetThrLysValHisGlyLysArgTyrAlaTyrLysPhe--- 362
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 1464 TATTACGACAAAACATCATCCACAGACAGCGGAGAAAGCTACGTGATCCGCTTGTG 1523
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 363 ---AspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlu 377
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 1524 TGTGACCTG-----CAGAGCTGCTGGGATCACCCCTGAG 1559
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

```

## RESULT 8

US-09-344-579-1  
Sequence 1, Application US/09344579

```

; Patent No. 6054316
; GENERAL INFORMATION:
; APPLICANT: Brenda F. Baker
; APPLICANT: Lex M. Cowert
; TITLE OF INVENTION: ANTISENSE MODULATION OF ETS-2 EXPRESSION
; FILE REFERENCE: PUS-0063
; CURRENT APPLICATION NUMBER: US/09/344, 579
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (291)..(1700)
US-09-344-579-1

```

Alignment Scores:

Pred. No.:	8.63e-37	Length:	2268
Score:	445.50	Matches:	127
Percent Similarity:	40.96%	Conservative:	43
Best Local Similarity:	30.60%	Mismatches:	105
Query Match:	18.23%	Indels:	141
	3	Gaps:	14

US-09-902-772-2 (1-451) x US-09-344-579-1 (1-2268)

```

QY 122 GlnArgArgValillevalprolaasprothleutrpserthrasphisValArgln 141
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 570 CACGGCGGCTGGGCAATCCAAAGAACCCCTGGCTGTGAGTAGACAAACAGGTATGCCAG 629
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 142 TrpLeuGlnTrpAlaVallylsglulrlyGlyleuProaspValaspileuleuPheGln 161
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 630 TGGCTTCTCGGGCCACCAATGAGTTCAGTGTGAGTGAACCTGATTCGACAGAGTTTC--- 666
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 162 AsnIleaspllylsglulrleucyslysmethrlyaspaasphelinarGluThrPro 181
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 687 GGCATGAATGGCAGATGCTGTGTAACCTTGGCAAGAGACGTTTCTGGAGCTGGGACCT 746
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 182 SerTyrAsnAlaaspileuleuSerHisleuHstlyrleuArgIleuArgIleThr 201
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 747 GACTTGTGGGTACATCTCTGAGAACATCTGAGCAAAATGATCAAGAAACCA--- 803
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 202 PheIlePheProasnThrSerValTyrProGlnAlaThrGlnArgIleThrArgPro 221
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 804 -----GAAAGACAGAAATCAATNGAAGAAATTCACACCTCAGCTGCTTCC 854
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

QY 222 -----AspLeuProTyr-GluGlnAl 228
    ||| ||| ||| :|||: ||||| ||||| |||||: |||
Db 855 CATTGATTAAACAGCAATACATTAGTTTGGCACAGAGAGCGCCCTTGAATATGACAG 914
    ||| ||| ||| :|||: ||||| ||||| |||||: |||

```





```

Db          427 TCACGTACAAAGTTCATTTCACAAACATGCTGTGTCATATACCAATTCATTGATG 468
Oy          366 -----GlytIleAlaGlnAlaLeuGlnProHisProGlnSerMet 381
Db          487 GCTTGCGTGGGGGTGCAGCCCGAGAGTGGCCCCCGCAGTCCCGGGGTGGCCACT 546
Oy          381 ytlstlyrProSerArbLeuProTyrMetSerSetTyrHisAlaHisProGlnMet 401
Db          547 TCGGTCCTCCCTCCCTCACGCC-----TCCG 573
Oy          401 snheValAlaPro-----HisProAlaLeuProValrhesers 416
Db          574 AGGCGGTGTCGCCCGACGAGACCCCGCTACACGACGCGCTTCATCTTCATCTT 633
Oy          416 erPhepheAlaAla 420
Db          634 CCTCTCTCTCGGCT 647

RESULT 10
US-09-021-715-1
; Sequence 1, Application US/09021715
; Patent No. 6194547
; GENERAL INFORMATION:
; APPLICANT: MayoThalassitis, George J.
;           Blair, Donald G.
;           Fisher, Robert J.
;           Beal Jr., Gregory J.
;           Athanasiou, Meropi A.
;           Sgouras, Dionyssios N.
; TITLE OF INVENTION: The ERF Genetic Locus and Its Products
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3634
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,715
; FILING DATE: 10-Feb-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Mackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 015280-229000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2667 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 123..1769
; OTHER INFORMATION: /note="human ERF (ETS2 Repressor
; Factor) cDNA"
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-021-715-1

Alignment Scores:
Pred. No.: 7 49e-26 Length: 2667
Score: 342.50 Matches: 84
Percent Similarity: 53.408 Conservative: 26

```

```

Best Local Similarity: 40.78%  Mismatches: 59
Query Match: 14.01%  Indels: 38
DB: 4  Gaps: 5

US-09-902-772-2 (1-451) x US-09-021-715-1 (1-2667)

QY 239 ProthrgInserLysAlahrcInProSerSerThValProLysrThrgInuasp--G 258
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 74 CCGGAGCCGCGCGCCCGGAAATCGGGGCGCTGGCCCCCGGGCCCCCGACGTGAAGACCC 133
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 258 InarProGInleuAaspProTyrgInlleuGlyProThrSerSerArgLeuAlaasp 278
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 134 GCGGACACAGGGGTTTGCTTCCCGGAT--GGGCTACA--ACCGAG-TCGTCC 166
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 278 roGLySerGLyInlleGInleuTPrgInPheLeuLeuGluLeuSerAspSerSera 298
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 187 CTGGCTCAAGGACGATCCAGCTGTGGCACTTATTCCTGGAGCTGCGGAGAGAGT 248
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 288 snSerAsnCySlleThrTPrgInuGLyThrAsnGLyLuphElysmetThrAspProasp 318
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 247 ACCAGGCGCTATTCCTGGCAGGGGAGCTACGGGGAATTCGTATCAAGACCTGATG 306
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 318 luValAlaArgrTgTgGLyGluArgLysSerLysProAsnmetAsnTyAspLysLeus 338
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 307 AGTGCGCCCGCTGTGGGCGCTTCGCAATGCAGAGCCCGCAGATGATACGACAGCTGA 366
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 338 eArGAlaLeuArgrTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 358
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 367 GCGGGCCCTGGCTATTAATCAACACGCAATTCGACACAGGACCAAGGGGAAAGGT 426
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 358 yTAlaTyLysrLysheAspPheHis----- 365
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 427 TCACCTCAAGTTCATTCATTCACAACTGTGTGTGTCATTACCATTCATTCATTCATG 486
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 366 -----GlylleAlaGlnAlaLeuGlnProHisProGluSerSermet 381
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 487 GGTGGCTGGGGGTGCAGTGCACAGAGTCCCGCCGACAGTCCGTGGGGGTAGCCACT 546
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 381 yTLYsTyTyProSerAspLeuProTyTyMetSerSerTyHisAlaHisProGlnLysMet 401
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 547 TCCGCTCCCTCCCTCCCAAGGCC-----TCCG 573
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 401 snPheValAlaPro-----HisProAlaLeuProValThrSerSers 416
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 574 AGGTCTGTCCCCCAGAGAGACCCCGCTCACACAGCCTGCTTCATTCATTCATCTT 633
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 416 erPhePheAlaAla 420
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 634 CCTCTCTCGGCT 647
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 11
US-08-875-944B-1
; Sequence 1, Application US/0887594AB
; Patent No. 6036542
; GENERAL INFORMATION:
; APPLICANT: FUJINAGA, Kei
; APPLICANT: YOSHIDA, Koichi
; APPLICANT: HIGASHINO, Fumihito
; TITLE OF INVENTION: CANCER CONTROL
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```









```
Db 922 -----CATCTCTCAGGAGGCGGAGCCGGAACCTTCTCCAGCCCCCTATGACACACC 972
QY 208 Ser-----ValTyrProGluAlaThrGlnArg----- 216
Db 973 AACGTGGAGCCCTGCCACCCCTACCCCGCAGAACTTCAACAGAGTACCATGACC 1032
QY 217 ---lIeThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAla----- 232
Db 1033 CCCTGTACGAACAGCTGGCCAGCCGCTT---CAAGCCAGGCTGGGGTCACTGGGCACA 1089
QY 233 -----TTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln 246
Db 1090 GGTACCCAGGCGCGGGGTGGTGAATCAAAACAGAGCCACAGACTTCGCTACGACTCAG 1149
QY 247 ProSerSerSerThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGln 266
Db 1150 ATGTCCCTGGATGTGCATCATGATGATGATCCACCCAGAGGGCTTCTGTGACCTCTCCAG 1209
QY 267 lIe----- 267
Db 1210 GTGATGGAGTGTGATTATGCTATGAAAAATCCCTTCGACATTCACAGATGATGCT 1269
QY 268 -----LeuGlyProThrSerSerArg-----LeuAlaAsn 277
Db 1270 GCATTGTCCCTAAAAAATTGTAAGAGAGACATCAAGCAGAGAGGATTGAGACTTTCGGG 1329
QY 278 ProGly-----SerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeu 292
Db 1330 AG-GGGCCACCCCTACAGCGCGGGGTGCTTACACTGTGCACTTGTGGCCCTG 1388
QY 293 LeuSerAspSerSerAsnSerAsnCysIlleThrTrpGluGlnThrAsnGlyGluPheLys 312
Db 1389 CTGGATGACCCCAACAATGCTCATTTCTGCTTGACAGCGCGGGGATGAGATTAA 1448
QY 313 MetThrAspProAspGluValAlaAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332
Db 1449 CTAATTTGAACCTGAAGAGGTGCCGCTCGGGGTATCCAGAAGAACCGCCAGCCATG 1508
QY 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352
Db 1509 AATTATGACAAAGCTGACCGCTCGCTCGATACTATATATGAAAGGATCATCATCAGAAAG 1568
QY 353 ValHisGlyLysArgTyrAlaTyrLysPhe 362
Db 1569 GTGGCTGGCAGACGCTACGTACAACTT 1598
```

Search completed: July 28, 2003, 07:41:33  
Job time : 76.6835 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 06:34:34 ; Search time 284.484 Seconds

(without alignments) : 3270.525 Million cell updates/sec

Title: US-09-902-772-2

Perfect score: 2444

Sequence: 1 MASTIKKALSVSEDSQSLFE.....IYPNTRLPAAHMPHGLCTYY 451

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1439767 segs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame-p2n -DEV-xlh  
-O-/cgn2\_1/USPTO.spool/US09902772/funat\_23072003.093658.14962/app-query.fasta\_1.1294  
-DB-published.Applications\_NA -OFMT-fastap -SUFFIX-rnpb -MINMATCH-0.1  
-LOOEXT-0 -LOOEXT-0 -UNITS-bits -START-1 -MATRIX-blosum62  
-TRANS-human40.cdi -LIST-45 -DOCALIGN-200 -THR SCORE-pct -THR MAX-100  
-THR MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MTLEN-0  
-MAXLEN-2000000000 -USER-US09902772.ecgn\_1.1.82.funat\_23072003.093658.14962  
-NCPU-6 -ICPU-3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES-0 -WAIT -DSELOCK-100  
-LONLOG -DEV-TIMEOUT-120 -WARN-TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5  
-FGAPOP-6 -FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : Published Applications\_NA:\*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PTI_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq3:*
13:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2444	100.0	1447	11	US-09-902-772-1
					Sequence 1, Appl1

2	2414.5	98.8	1528	11	US-09-902-772-3
3	2182	89.3	3166	15	US-10-205-823-98
4	948	38.8	567	10	US-09-864-761-18410
5	863	35.3	473	10	US-09-864-761-20472
6	712	29.1	549	10	US-09-923-779-54
7	650	26.6	420	10	US-09-864-761-17305
8	650	26.6	454	10	US-09-864-761-1653
9	638	26.1	467	12	US-09-918-995-23366
10	511	20.9	1752	10	US-09-850-799-1
11	445.5	18.2	2265	11	US-09-954-531-955
12	444.5	18.2	2188	11	US-09-920-300A-1716
13	444.5	18.2	2188	15	US-10-033-528-1716
14	444.5	18.2	2268	11	US-09-920-300A-1693
15	444.5	18.2	2268	15	US-10-033-528-1693
16	444.5	18.2	3692	15	US-10-106-698-824
17	431.5	17.7	1884	11	US-09-925-300-820
18	407	16.7	2532	15	US-10-037-270-869
19	374	15.3	553	10	US-09-864-761-7427
20	371	15.2	225	10	US-09-864-761-21761
21	371	15.2	348	10	US-09-864-761-29364
22	371	15.2	477	10	US-09-864-761-5035
23	371	15.2	519	10	US-09-864-761-12798
24	339	13.9	3178	15	US-10-108-605-124
25	339	13.9	3178	15	US-10-108-605-128
26	331	13.5	473	12	US-09-918-995-27904
27	322	13.2	2064	12	US-09-884-363-3
28	322	13.2	2064	15	US-10-167-485-1
29	322	13.2	2333	11	US-09-920-300A-1788
30	322	13.2	2333	11	US-09-880-107-3316
31	322	13.2	2333	15	US-10-033-528-1788
32	306	12.5	2410	12	US-09-884-363-1
33	303	12.4	416	12	US-09-918-995-8472
34	301	12.3	1894	10	US-09-126-9458-1
35	301	12.3	1905	10	US-09-866-356-2
36	301	12.3	1942	15	US-10-157-931-83
37	300.5	12.3	2952	10	US-09-925-297-257
38	300.5	12.3	4106	15	US-10-084-817-187
39	296	12.1	300	15	US-10-177-063-15
40	287	11.7	2256	15	US-10-098-841-317
41	282	11.5	1087	15	US-10-102-806-215
42	281	11.5	300	15	US-10-177-063-16
43	280	11.5	1894	10	US-09-841-963A-1
44	279	11.4	165	10	US-09-864-761-23294
c	260.5	10.7	1429	11	US-09-764-864-320

#### ALIGNMENTS

RESULT 1  
US-09-902-772-1  
Sequence 1, Application US/09902772  
Patent No. US20020164739A1  
GENERAL INFORMATION:  
APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et  
TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
TITLE OF INVENTION: the Proteins  
FILE REFERENCE: Chugai selyaku kabushiki kaisha 5001  
CURRENT APPLICATION NUMBER: US/09/902,772  
PRIOR FILING DATE: 2001-07-12  
CURRENT FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: US/08/878,177  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 1447  
TYPE: DNA  
ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA  
US-09-902-772-1

Alignment Scores:  
Pred. No.: 1.09e-256 Length: 1447  
Score: 2444.00 Matches: 451  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.008 Mismatches: 0  
 Query Match: 100.008 Indels: 0  
 DB: 11 Gaps: 0

US-09-902-772-2 (1-451) x US-09-902-772-1 (1-1447)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 DB 63 ATGGCAAGCACTATTAGAGAGCATTTATCAGTGTGATGAGAACACAGCTCTTTGAG 122  
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 DB 123 TGTGCTACGAGATGCCCGACCTTGGCAAGACAAATGACAGCTCTCTCCAGTGA 182  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 DB 183 TATGGCAAAACATCAAGATGAGCCCGCGGTTCACGACGAGCTGTATACACCCC 242  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 DB 243 CCGGCCAGAGTTACATTAGATGAGAGTAAACCAACAGGTTAATGGTCAAGGAAT 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspVal 100  
 DB 303 TCACCTGATACGACGCGTGGCAAGAGGAAATGGTTAGCACTTCAGACAAATGTT 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluHisIleProProProAsnMetThrThr 120  
 DB 363 GGGATGAACATATGGAACCTACATGGAAGAGAGCATATCCGCTCCAAATATGCAACC 422  
 QY 121 AsnGluArgValIleValProAlaAspProThrIleutrpSerThrAspHisValArg 140  
 DB 423 AATGAAAGAGAGTATTTGCGACAGATCTACGTTATGAGACACAGCATGTACGG 482  
 QY 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160  
 DB 483 CAGTGGCTGAGTGGGAGAGTGAAGAGATGCTTCCACACGCGGACATCTTGTTC 542  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 DB 543 CAGAACATGTATGGGAAAGAGTGTGTAAATGACCAAGATGACTCCAGAGACTCAGC 602  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAla 200  
 DB 603 CCAGAGCTATACGACAGATATCCCTGTCACACCTACACCTCAAGAGAGAGAGCC 662  
 QY 201 ThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArg 220  
 DB 663 ACTTTATTTTCCAAATATACATCATGTTATCCAGAGACACGCAAAATATACACAGG 722  
 QY 221 ProAspLeuProTyrGlnGlnAlaLysArgSerAlaTrpHisSerHisSerProThr 240  
 DB 723 CCAGATTATCTTATGAGCAAGAGAGATACAGCGAGATCCAGCATCCCACT 782  
 QY 241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgPro 260  
 DB 783 CAGTCAAAAGCTACCCACCATCATCTTCAACAGTGGCCAAACAGAGACAGCGCTCT 842  
 QY 261 GlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280  
 DB 843 CAGTGTAGATCTTATGATTCCTTGGACGACACAGACGCGCTGTGCAAAATCCAGGAGT 902  
 QY 281 GlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuSerAspSerSerAsnSerAsn 300  
 DB 903 GGGCAGATACAGCTATGCGAGTCTCACTGAGACTTCTGTGCGACACCTCCAACTCCAAC 962  
 QY 301 CysIleIleThrTrpGluGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAla 320  
 DB 963 TGCATACCTCTGGAGGGCAAAATGGGAGTCAAGATGACAGACCTGTATGAAGTGGCT 1022  
 QY 321 ArgArgTrpGlyLysGlyLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAla 340  
 DB 1023 CGGCGTTGGGAGAGAGAAACCAACCTAACATGACTATGACAAACCTCAGCCGTGCA 1082

QY 341 LeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360  
 DB 1083 CTTCGCTACTACTATGACAAAAATATATATGACTTAAGTTCATGATTAACGCTATGCTTAC 1142  
 QY 361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerSerMet 380  
 DB 1143 AAATTTATTTCCAGGAAATCGCTCAGGCCCTCCAGCTCAACCTCCAGATATCATCATG 1202  
 QY 381 TyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMet 400  
 DB 1203 TACAATACCATATAGACCTCCCTACATGATGTTCTTACCATGACACCCCCAGAAATG 1262  
 QY 401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPheAlaAla 420  
 DB 1263 AACTGTAGCTCCCATCCCGCTGGTGGCGGTAACTCATCCACACTTTTGTGCTGCC 1322  
 QY 421 ProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrProAsnThrArgLeuProAla 440  
 DB 1323 CCTAATCATATCTGGAATTCACCAACTGAGGCGATCTACCCCAATACAGGCTGCCAGCT 1382  
 QY 441 AlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 DB 1383 GCTCATATGCTTCCATTTTGGCACCCTACTAC 1415

RESULT 2  
 US-09-902-772-3  
 : Sequence 3, Application US/09902772  
 : Patent No. US20020164739A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et  
 : TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
 : FILE REFERENCE: chugai selyaku kabushiki kaisha 5001  
 : CURRENT APPLICATION NUMBER: US/09/902,772  
 : PRIOR FILING DATE: 2001-07-12  
 : PRIOR APPLICATION NUMBER: US/08/878,177  
 : NUMBER OF SEQ ID NOS: 7  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO: 3  
 : LENGTH: 1528  
 : TYPE: DNA  
 : ORGANISM: c-erg gene, chicken DNA  
 : US-09-902-772-3

Alignment Scores:  
 Pred. No.: 1,94e-253 Length: 1528  
 Score: 2414.50 Matches: 450  
 Percent Similarity: 94.14% Conservative: 0  
 Best Local Similarity: 94.14% Mismatches: 1  
 Query Match: 98.79% Indels: 27  
 DB: 11 Gaps: 1

US-09-902-772-2 (1-451) x US-09-902-772-3 (1-1528)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 DB 63 ATGGCAAGCACTATTAGAGAGCATTTATCAGTGTGATGAGAACACAGCTCTTTGAG 122  
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 DB 123 TGTGCTACGAGATGCCCGACCTTGGCAAGACAAATGACAGCTCTCTCCAGTGA 182  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 DB 183 TATGGCAAAACATCAAGATGAGCCCGCGGTTCACGACGAGCTGTATACACCCC 242  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 DB 243 CCGGCCAGAGTTACATTAGATGAGAGTAAACCAACAGGTTAATGGTCAAGGAAT 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspVal 100



```

Db      303 TCACCTGATGACTGCAGCGGTGGCAAAAGAGGAAATAGTTAGCAGATTACAGCAATGTT 362
Qy      101 GYMETASNTYRGLYSERTYMETGLUGLYUSHSISLEIROPPTROASMETHTHR 120
Db      363 GGGATGAATGAGGCTACCTAGTGAAGAAGCAATATCCGCTCCAAATATGCAAC 422
Qy      121 ASNGUATARGVALILEVALPROALAASPROTHLEUTPSETHTASPHISVALRY 140
Db      423 AATGACGAAGAGTATTGTCTGCACACATCTTACGTTATGAGACACGACCATATGAG 482
Qy      141 GINTPLEUGLUTRPALEVALYSGIUTYRGLYLEUPROASPVALAASPILEULEUPHE 160
Db      483 CAGTGGCTGAGAGTGGCAGAGTGAAGAGATGAGTCTTCACAGAGTGCACATCTGTTGTC 542
Qy      161 GINANILEASPGLYLYSGILEUCYSLSYMETHTHRYSASPSPHGLINARGLIETHR 180
Db      543 CAGAACATTGATGGGAAAGAGTGTGTAAATGACCAAAAGTACTTCAGAGACTCAG 602
Qy      181 PROSEPTYRASNALASPILEULEUSERHISLEUHSITYRLEUARGLUARG----- 198
Db      603 CCGAGCTATTAAGCAGATATCTCTCTGTACACACTACACTACCTTAGAGAGACTCTCT 662
Qy      198 ----- 198
Db      663 CCACATTTGACTTCAGATGATGTTGATAGGCGCTTACAAAACCTCCACGGTTATGCA 722
Qy      199 -----GLYALATHRPHLEPHEPROASNTHTSERVALTYRPROGLUALA 213
Db      723 GCTAGAAAACACAGAGAGCCACCTTTATTTTTCAAATACATCAGTTTACCAGAACCA 782
Qy      214 THRGINARGILETHRTHRARGPROASPLEUPROTYRGLUGLINALARGSERALATRP 233
Db      783 ACGCAAAATATACACAGGCCAGATTACCTTATGAGCAAGCAGAGAGATCAGCGTGG 842
Qy      234 THRSERHISERHISPROTHGLINSELYSALAETHRGLNPROSETHSERTHRYALPRO 253
Db      843 ACGAGTTCACAGCCATCCACTCAGTCAAGCAAAAGCTACCCACATCATCTTCAACAGTGGCC 902
Qy      254 LYSHTHRLUASPGINARGPROGLINLEUASPROTYRGLINILEUGLYPROTHSER 273
Db      903 AAAACAGAGACAGCGCTCTCAGTATGATCTTATGAGATTTGGACCCAGCAGCAGC 962
Qy      274 ARGLEUALAASNPGLYSERGLYGLINILEGINTLEUTPGLINPHELEULEUGLEU 293
Db      963 CGTCTTGAATTCAGAGGAGTGGCAGATACAGCTATGCGAGTTCCTACTGAGGCTTCTG 1022
Qy      294 SERASPSERASNSERASNCYSILETHRTPLUGLINTHRASNGLYGLUPHELYSMET 313
Db      1023 TCGGACAGCTCCCACTCACTGATCATCTGGGAGGCGCAAAATGGGAGATTCAAGATG 1082
Qy      314 THRSAPROASPGILUALAARGARGTRPGLYGLUARGLYSERLYSERSPROASMETASN 333
Db      1083 ACAAGCCCTGATGAAAGTGGCTGGCGTTGGGAGAGAGGAAACCAAACTAACATGAC 1142
Qy      334 TYRASPLYLSERASERARGALALEUARGTYRTYRTYRASPLYASNIEMETHLYSVAL 353
Db      1143 TATACAAACTACACCGCTGCGCTACTACTATGACAAATAATATTATACATGAACTT 1202
Qy      354 HISGLYLSARGTYRATYRLYSPHASPHEHISGLYILEALAGLINALALEUGLIPRO 373
Db      1203 CATGCTAAACGCTATGCTACAAATTTGATTCCACGGAATCGCTCAGCGCTCAGCGCT 1262
Qy      374 HISPROPGIUSERSEMETTYRLYSTYRPROSERASPLEUPROTYRMETSERSETYR 393
Db      1263 CACCTCCAGATATCCATGATACAAATACCATCAGACCTCCCTCATGATGTTCTTAC 1322
Qy      394 HISLAHSPROGLINLYSMETASNPHEVALALAPROHISPPROPROALALEUPROVALTHR 413
Db      1323 CATGCACACCCCGAAGATACACTTTGTAGCTCCCATCCCTGCTTGGCCCGTAAACC 1382
Qy      414 SERSETHRPHLEALALAPROASNPROTYRTPASNSERPROTHRGLYGLYILETYR 433

```

```

Db      1383 TCATCCAGCTTTTGTCTGCCCTTAATCCATACCTGGAATTCACCAAGTGGACATCTAC 1442
Qy      434 PROASNTHTARGLEUPROALALAHISMETPROSETHISLEUGLYTHRTYR 451
Db      1443 CCCAATACCAGGCTCCAGCTGCTCATATGCTTCCCATCTTGGCAGCCTACTAC 1496

RESULT 3
US-10-205-823-98
; Sequence 98, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gordatcheva, Bella
; APPLICANT: Hoerscht, Sebastian
; APPLICANT: Kamalakar, Shubhangt
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatf, Karen
; APPLICANT: Zhao, Xumel
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-98.

Alignment Scores:
Pred. No.: 1,27e-227 Length: 3166
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 88.67% Mismatches: 24
Query Match: 89.28% Indels: 14
Gaps: 4

US-09-902-772-2 (1-451) x US-10-205-823-98 (1-3166)
Qy      2 ALASERTHRLIYLSGUALALEUSERVALVALSERGLUASPGINSETHLEUPHEGLUCYS 21
Db      281 GCACCTCATATACAGGAAGCCTTATCAGTTGTGAGTGAGACCAAGTCTGTGAGAGT 340
Qy      22 ALATYRGLYSERPROHISLEUALALYSHTRGLUMETHRALASERSETSERSETGLUTYR 41
Db      341 GCCTACGGAAGCCACACCTGGCTPAAGACAGATGACCGGCTCTCTCCAGCGACTAT 400
Qy      42 GLYGINTHRSERLYSMETSERPROARGVALPROGLINLASPTPLUSERGLIPROPRO 61
Db      401 GGACAGACTTCCAGAGTGAAGGCCACGGCTCCTCAGCAGATTTGGCTGTCTCAACCCCA 460
Qy      62 ALAARGVALTHRLIYLSMETGLUCYSASNPROASNGLINALASNGLYSERARGASNSER 81
Db      461 GCCAGGGTCAACCAATGAATGTAACCTTGACCAAGTGGAATGCTCAAGGAACCTCT 520
Qy      82 PROASPSAPSYSSERVALALALYSGLYGLYLSMETVALSERSETHSERASPSVALGLY 101

```

D	521	CCTGATGATGACAGGTGGCCAAAGCGGGAAAGATGTTGGGACGCCACACCGTTGGG	580
Q	102	MetasnrYrgLYSerTYmetcIuGlulYshisIleProProAsnMethrThrAsn	121
D	581	ATGAACACTACGGCAGCTACATGAGGAGAAAGCAATGCCAACCCCCAAACATGCACAGAAC	640
Q	122	GIuArGaArvAlIleValProIaIasProMhrIeuthrPserThrAspHisValrgrIn	141
D	641	GAGCCCAAGATTATCTGCCACACCAATCCCTTCGCTATGAGTACAGACCATCTGGCGGAC	700
Q	142	TrPleuGIuTrPrAlaValIaVallYsgIuVrYgrGlyLeuProAspValAspIleuLeuPheGln	161
D	701	TGGCTGGAGTGGGGCGGTGAAGAATATGGCTTCGACAGCTCAACATCTTGTATTCCAG	760
Q	162	AsnIleaspgLYIysGLuIeucYsLYsmethrLYsasPspPheGlnarGLeuthrPro	181
D	761	AACATCGATGGGAAGAACTGTGCAAGATGACCAAGACAGACTTCACAGAGCTCACCCCC	820
Q	182	SerTYrAsnAlaAspIleleuLeuSerHisIleuHisTYrLeuArGluArGlyAlaThr	201
D	821	AGCTAACACCGCGACATCTCTTCACATCTCCACTCCACCTCCAGAGAAGCTCTCTCCA	880
Q	202	PheIlePheProAsnThrSerValTYrProGluAlaThrGlnArGIIeThrThrArpPro	221
D	881	CATTGG-----ACTTCA-----GATGATGTTGATTAAGCCTTACAAACTCTCCA	925
Q	222	-----AspleuProTYrGluGlnAlaArGArSerAlaTrp	233
D	926	CGGTTAATGACGATCTAAGAAACACAGATTATACATATGAGCCCCCAGAGATCACCGTGG	985
Q	234	ThrSerHisSerHisProThr---GlnSerLYsAlaThrGlnProSerSerSerThrVal	252
D	986	ACCGGCTACGGCCACCCACCGCCACGAGTGGAAAGCTGCTCAACCATCTCTCCACAGTG	1045
Q	253	ProLYsThrGluAspGlnArpProGlnIeuaSPProTYrGlnIleleuGlyProhrSer	272
D	1046	CCCAAAATCTGAAGACCAAGCGTCTAGTTAGATCTTATCAGATTCTTGGACCAACAAGT	1107
Q	273	SerArGleuAlaAsnProGlySerGlyGlnIleGlnIeuthrPglInPheLeuDeuGluIeu	292
D	1106	AGCGCCTTGCMAATCCAGAGCAGTGGCCAGATCCAGCTTGGCAGTCTCTCTGGAGTGC	1165
Q	293	LeuSerAspSerSerAsnSerAsnCYsIleThrTrpGluGlyThrAsnGlyIuPheLYs	312
D	1166	CTGTGCGACAGCTCCACACTCGCATCCACTCTGGGAAGGCCAACCAAGGGGAGTTTCAAG	1225
Q	313	MetThrAspProAspGluValAlaIaArGrArTrpGluIuArGlySerLYsProAsnMet	332
D	1226	ATGACGGAATCCCGAGAGAGGTGGCCCCGGCGCTGGGGAGAGCGGAAGCAACCAACATG	1285
Q	333	AsnTYrAspLYsLeuSerArGAlaIeuaArGTYrTYrTYrAspLYsAsnIleMethrLYs	352
D	1286	AACATCGAATAGCTAGCGCGCGCCCTCGTTACTACTATGACAAACATCATGTGACCAAG	1345
Q	353	ValHisGLyLYsArGTYrAlaTYrLYsPheAspPheHisGLyIleAlaGlnAlaLeuGln	372
D	1346	GTCCTATGGGAAGCGGTACGCTACAAAGTTTCGACTTCCACAGGATGGCCCAAGGCTCCAG	1405
Q	373	ProHisProProGluSerSerMetLYrLYsTYrProSerAspLeuProTYrMetSerSer	392
D	1406	CCCCACCCCCCGGAATCTACTCTGTACAAAGTACCCCTAGACCTCCCTACATTGGGCTCC	1465
Q	393	TYrHisAlaHisProGlnLYsMetAsnPheValAlaProHisProProAlaIeuthrProAl	412
D	1466	TATTCACGCCCAACCCACAGAGTGAATTTGTGGCGGCCCAACCTTCACACCTCCCGG	1522
Q	413	ThrSerSerSerPhePheAlaAlaProAsnProTYrTrpAsnSerProThrGlyLYsIle	432
D	1526	ACATGTTCCAGTTTTTTTGTGGCCCCCAACCATATCTGGAAATTCACCAACTGGGGGTATA	1585
Q	433	TYrProAsnThrArGLeuProAlaAlaHisMetProSerHisLeuGlyThrTYrTYr	451

```

DB
15861ACCCCACTAGGCTCCACAGCAGCATATGCTTCTCATCTGGGCACTTACTAC 1642

RESULT 4
US-09-864-761-18410/C
; Sequence 18410, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18410
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000163.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
; OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 9.00e-99
; OTHER INFORMATION: EST_HUMAN HIT: R87572.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: M17254.1, EVALUE 0.00e+00

```



```
Db 233 GCGTACAGTTCGACCTCCAGGATCGCCAGGCGCCCTCCAGCCCGCCGAGTCA 174
Qy 379 SerMetYrlySyrProSerAspLeuProTyrMetSerTyrYHsAlaHisProIn 398
Db 173 TCTGTGACAGTACCCCTCCAGACCTCCGATACATGGCTCTCTATCAGCCCGACAG 114
Qy 399 LysMetAspHeValAlaProHisProAlaLeuProValThrSerSerPhePhe 418
Db 113 AAGATGAACCTTGTGGGCGCCCGCCCTCCAGCCCTCCCGATCATCTTCAGTTT 54
Qy 419 AlaAlaProAsnProTyrTTPAsnSerProThrGlyGlyIleTyrProAsn 435
Db 53 GCTGCCCAACCACTACTGGAATTCACCAACTGGGGGTATATACCCCAAC 3

RESULT 6
US-09-923-779-54
: Sequence 54, Application US/09923779
: Patent No. US20020076721A1
: GENERAL INFORMATION:
: APPLICANT: Pyle, Ruth A.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Kalos, Michael D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.553
: CURRENT APPLICATION NUMBER: US/09/923,779
: CURRENT FILING DATE: 2001-08-06
: NUMBER OF SEQ ID NOS: 155
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 54
: LENGTH: 549
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 356, 363, 417, 433, 442, 461, 463, 464, 469, 479, 485, 489,
: LOCATION: 537, 545
: OTHER INFORMATION: n = A,T,C or G
US-09-923-779-54

Alignment Scores:
Pred. NO.: 2,396-68 Length: 549
Score: 712.00 Matches: 128
Percent Similarity: 95.59% Conservative: 2
Best Local Similarity: 94.12% Mismatches: 5
Query Match: 29.13% Indels: 1
DB: 10 Gaps: 0

US-09-902-772-2 (1-451) x US-09-923-779-54 (1-549)
Qy 317 AspGluValAlaArgATGTPGlyGluArgLysSerLysProAsnMetAsnTyrAspLys 336
Db 2 GAGCAGGTGGCGCGCGCTGGGGAGAGCGGAAGCAACCCACATGAGACTAGATAG 61
Qy 337 LeuSerArgAlaLeuAlaGlyTyrTyrAspLysAsnIleMetThrLysValHisGlyLys 356
Db 62 CTCGCGCGCGCGCTCCCTTACTACTATGACAAAGCATCATGACCAAGTCCATGGGAG 121
Qy 357 ArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProPro 376
Db 122 CGCTACGCTTACAAAGTTCGATTCACGGGATCGCCCAAGGCGCTCCAGCCCGCCG 181
Qy 377 GluSerSerMetYrlySyrProSerAspLeuProTyrMetSerTyrYHsAlaHis 396
Db 182 GAGTCATCTCTGTACAGTACCCCTCCAGACCTCCCGATACATGGGCTCTATCAGCCCGAC 241
Qy 397 ProGlnLysMetAspHeValAlaProHisProAlaLeuProValThrSerSerSer 416
Db 242 CCACAGGAAGATGAACCTTGTGGGCGCCCGCCCTCCAGCCCTCCCGATCATCTTCAGT 301
Qy 417 PheHeValAlaProAsnProTyrTTPAsnSerProThrGlyGlyIleTyrProAsnThr 436
```

```
Db 302 TTTTGTGCTGCCCAACCACTGCAATTCACCAACTGGGGGTATATACCCNTCACT 361
Qy 437 ArgLeuProAlaAlaHisMetProSerHisLeu-GlyThrTyrTyr 451
Db 362 ANGCTCCCGACAGACCATATGCTTTTCATCTGTGGGCGACCTTACTAC 407

RESULT 7
US-09-864-761-3705/c
: Sequence 3705, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DEIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
: SEQ ID NO 3705
: LENGTH: 420
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AP000021.2, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN B474, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
US-09-864-761-3705
```

## Alignment Scores:

Pred. No.:	9,466-62	Length:	420
Score:	650.00	Matches:	114
Percent Similarity:	98.31%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	2
Query Match:	26.60%	Indels:	0
DB:	10	Gaps:	0

US-09-902-772-2 (1-451) x US-09-864-761-3705 (1-420)

```

OY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrAspLysAsnIleMetThrLysVal 353
    |||||||
DB 418 TACGATAGCTCAGCGCGCCCTCCGTTACTACTATGACAGAACATCATGACCAAGTTC 359
    |||||||
OY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
    |||||||
DB 358 CATGGGAAGCGCTACGCTACAGAGTTCCACTTCCACGGGAGTCGCCAGGCCCTCCAGCCC 299
    |||||||
OY 374 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
    |||||||
DB 298 CACCCCGCGAGTCTCTCTGTACAGTACCCCTCAGACCTCCGTACATGGGCTCTAT 239
    |||||||
OY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
    |||||||
DB 238 CACGCCACCCACAGAAATGAACTTTGTGGGCCCCCAGCCCTCCAGCCCTCCGCTGACA 179
    |||||||
OY 414 SerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGlyGlyLeTyr 433
    |||||||
DB 178 TCTTCCAGTTTGTGCTGCCCCCAAAACCATACGGAATTCACCAACTGGGGGTATATAC 119
    |||||||
OY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
    |||||||
DB 118 CCCAACACTAGGCTCCGCCACCAAGCATATGCTTCTCATCTGGGCACTTACTAC 65
    |||||||

```

## RESULT 8

US-09-864-761-1653/c  
 ; Sequence 1653, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wensheng  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 FILE REFERENCE: Aecmica-X-1  
 CURRENT APPLICATION NUMBER: US/09/864,761  
 PRIOR FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263,6  
 PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 1653

LENGTH: 454  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AP000163.1  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4  
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95  
 US-09-864-761-1653

## Alignment Scores:

Pred. No.:	1,066-61	Length:	454
Score:	650.00	Matches:	114
Percent Similarity:	98.31%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	2
Query Match:	26.60%	Indels:	0
DB:	10	Gaps:	0

US-09-902-772-2 (1-451) x US-09-864-761-1653 (1-454)

```

OY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrAspLysAsnIleMetThrLysVal 353
    |||||||
DB 452 TACGATAGCTCAGCGCGCCCTCCGTTACTACTATGACAGAACATCATGACCAAGTTC 393
    |||||||
OY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
    |||||||
DB 392 CATGGGAAGCGCTACGCTACAGTGTCCACGCGGAGTCGCCAGGCCCTCCAGCCC 333
    |||||||
OY 374 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
    |||||||
DB 332 CACCCCGCGAGTCTCTGTACAGTACCCCTCAGACCTCCGTACATGGGCTCTAT 273
    |||||||
OY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
    |||||||
DB 272 CACGCCACCCACAGAAATGAACTTTGTGGGCCCCCAGCCCTCCAGCCCTCCGCTGACA 213
    |||||||
OY 414 SerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGlyGlyLeTyr 433
    |||||||
DB 212 TCTTCCAGTTTGTGCTGCCCCCAAAACCATACGGAATTCACCAACTGGGGGTATATAC 153
    |||||||
OY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
    |||||||
DB 152 CCCAACACTAGGCTCCGCCACCAAGCATATGCTTCTCATCTGGGCACTTACTAC 99
    |||||||

```

## RESULT 9

US-09-918-995-23356  
 ; Sequence 23356, Application US/09918995  
 ; Publication No. US20030073623A1  
 ; GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.  
 TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

[illegible]

```

Db      1047  GCTGTACCAACCCCGGGCTTGCAAGCCCCCTCCGGGGCCCTTTGGCGGGGTGGCGGGCCG 1106
QY      444  oSerHisIleuclyIhrTyIyr 451
          |||||
          ::|||
Db      1107  TTCGCACYTGGGGGGTCAWTAT 1128

RESULT 11
US-09-954-531-955
; Sequence 955, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 955
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-955

```

Alignment Scores:			
Pred. No.:	2,14e-38	Length:	2269
Score:	445.50	Matches:	127
Percent Similarity:	40.96%	Conservative:	43
Best Local Similarity:	30.60%	Mismatches:	105
Query Match:	18.23%	Indels:	141
DB:	11	Gaps:	14
US-09-902-772-2 (1-451) x US-09-954-531-955 (1-2269)			
QY	122	GLuARGATValLIeValProAlaSProThRIeuTPSerThrASpHisValArgIn	144
DB	571	CAAGGGCGCGTGGCCATTCACAAAGACCCTGGCTGGATGAGCAACAGTGTGGCAG	630
QY	142	TrPLeuGIuTPraValAlaYllySGluIuVrGLyeuPProASpValAspIleuLeuPheIn	161
DB	631	TGGCTTCTCTGGGCCACCAATAGATTCAGTCTGGGAGACGTGAATTCTGCAGAGTTC	687
QY	162	AsnIleAspGlyLysGluLeuCysLysMetThrIlySAspAspPheGlnArgLeuThrPro	181
DB	688	GGCATGAAGTGCACAGATGCTGTGTAACTTGGCAAGAACCTTTTCGAGCTGCACCT	747
QY	182	SeTrYrAsnAlaAspIleuLeuSerHisIeuHisTrIleuArgGluArgGlyAlaThr	201
DB	748	GACTTTGGTGGTACATTCCTCTGGGAAATCTGGAGCAAAAGATCAACAAAGAAACAA	804
QY	202	PheIlePheProAspThrSerValIyrProGluAlaThrArgIleThrThrArgPro	221
DB	805	-----GAAAGACAGAGATCAATATGAGAAATTCACACCTTCACCTCGGTCT	855
QY	222	-----AspLeuProTYr-GluGlnAl	228
DB	856	CATTGATTTACAGACATCATTTAGTTTGGCACAGACAGCGGCCCTATGCAATCAG	915
QY	228	aArgArg-----SerAlaTrpHiser-----HisSerHisProth	240
DB	916	ACACAGAAATTAACCAAGGCGGCTCTCGGACAGACATGTCTCGGCGCTCACACACCAGC	975

[illegible]

```

; LENGTH: 2188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1716

Alignment Scores:
Pred. No.: 2,61e-38 Length: 2188
Score: 444.50 Matches: 125
Percent Similarity: 41.28% Conservative: 43
Best Local Similarity: 30.71% Mismatches: 105
Query Match: 18.19% Indels: 135
DB: 11 Gaps: 13

US-09-902-772-2 (1-451) x US-09-920-300A-1716 (1-2188)

QY 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgln 141
   :::::::::::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 CAGGGGGCCGCGGCGCATTCACAAAGAACCCCGGCGTGTGAGTAGACAGCAACAGTATGCCAG 550

QY 142 TripleGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 TGGCTTCTGCGGCCACCAATGATTCAGTCTGTGTAACGTGAATCTGCAGAGGTTTC--- 607

QY 162 AsnIleAspGlyLysGluLeuCyLysMetThrLysAspAspPheGlnArgLeuThrPro 181
   :::::::::::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 GGCATGAAATGGCCAGATCTGTGTAACTTGGCAAGGAACGCTTCTGAGCTGCGACCT 667

QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
   ::|:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 GACTTTGTGGTGATCTCTCTGAGAACATCTGGACAAATGATCAAAAGAAACCAA--- 724

QY 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgPro 221
   ::|:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 -----GAAAGACAGAAAGATCATATGAAAGAAATTCACACCTCACCTCCGTTCCCT 775

QY 222 -----AspLeuProTyr-GluGlnAl 228
   ::|:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 776 CATTTGATTAACAGCAATACATTAGTTTGGCACAGACGCGCCCTATGGAAATCCAG 835

QY 228 AlaArgArg-----SerAlaTrpThrSer-----HisSerHisProth 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 836 ACACAGAAATTAACCCAAAGGGCGCTCTGAGACAGATGATGCCGCTCCACACCCAGC 895

QY 240 RginSerLysAlaThrGln-----ProSerSe 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 896 GTACTCAGCTCTGAGCAGGAGTTTCAGATGTTCCCAAGTCTCGCTCAGCTCCGTCAGC 955

QY 249 rSerThr----- 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 956 GTCACTACTGCTCTGTCACTGACGACTCCCGACAGCAACTTGAATTGGCTCACCAC 1015

QY 252 -----ValProLysThr-----GluAs 257
   ::|:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1016 AATTCTGGAGCTCCCAAGACAGACGACTCCCTGAGAAAGGTTGGGACAGCTTGCAGAGC 1075

QY 257 pGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro----- 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1076 TCAGACTCCCTCTCCAGTCTGTGAACAGCAGTCTGCTTGTGTGATGTGCACGGGTT 1135

QY 270 ----- 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1136 CCTTCTTCGAGAGCTTCGAAGATGACTGACGCCAGTCTCTGCTCAATTAAGCCAAACC 1195

QY 271 -----ThreSerArgLeuAlaAsnPro----- 278
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1196 ATGCTTTTCAAGATTACATCAAGAGAGG-AGTGAACCCAGTGGAGCAAGCAACACAGT 1254

QY 279 -----GlySerGlyGlnIleGlnLeuTrpGln 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1255 TATACCTGACAGCTGTGTGGCCGCTTCACAGGAAGGAGACATTAATTCAGTGTGGCAGTT 1314

QY 288 eLeuLeuGluLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpIleGluGlyThrAs 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1315 TCTCTGAGAGCTGCTATCAGACAAATCTCCAGTCAATTCATCACTGAGTGGAGACCG 1374
QY 308 nGlyIlePheLysMetThrAspProAspGluValAlaArgArgTrpGlyArgLysSe 328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1375 ATGGAGATTATAGCTCCGCCGACCCGATGAGGTGGCCGCGGTGGGAAAGAGAAATA 1434

QY 328 rLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAs 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1435 TAAGCCCAAGATGAAGTACAGAGAGAGCTGAGCCGGGGCTTACGCTACTATTCAGCAAGAA 1494

QY 348 nIleMetThrLysValHisGlyLysArgTyrAlaTyrLysPhe----- 362
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1495 CATATCCACAAGACGTGCGGGAAGCGCTACGTGATCCGCTGTGTCGACCTCCAGAA 1554

QY 363 -----AspPheHisGlyIleAlaGlnAlaLeuGlnProHisPr 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1555 CTTCGTGGGTTTCACGCCCGGAGAACTGACAGCCATCTG---GGCGTCCAGCCCGAAC 1611

QY 375 oProGluSerSerMetTyrLysTyrProSerAspLeuPro---TyrMetSerSerTyrH 394
   ::|:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1612 GGAGGACTGAGAGTGCAGCGGACACACCTGAGCCGCCAGGCTGCTGAGTGAAGTGGGA 1671

QY 394 sAlaHisProGlnLysMet 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1672 AGCCATCTGATCCAGCTG 1690

RESULT 13
US-10-033-528-1716
; Sequence 1716, Application US/10033528
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ. ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 1716
; LENGTH: 2188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-1716

Alignment Scores:
Pred. No.: 2,61e-38 Length: 2188
Score: 444.50 Matches: 125
Percent Similarity: 41.28% Conservative: 43
Best Local Similarity: 30.71% Mismatches: 105
Query Match: 18.19% Indels: 135
DB: 15 Gaps: 13

US-09-902-772-2 (1-451) x US-10-033-528-1716 (1-2188)

QY 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgln 141
   :::::::::::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 CAGGGGGCCGCGGCGCATTCACAAAGAACCCCGGCGTGTGAGTAGACAGCAACAGTATGCCAG 550

QY 142 TripleGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 TGGCTTCTGCGGCCACCAATGATTCAGTCTGTGTAACGTGAATCTGCAGAGGTTTC--- 607

QY 162 AsnIleAspGlyLysGluLeuCyLysMetThrLysAspAspPheGlnArgLeuThrPro 181
   :::::::::::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 GGCATGAAATGGCCAGATCTGTGTAACTTGGCAAGGAACGCTTCTGAGCTGCGACCT 667

QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
   ::|:: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 GACTTTGTGGTGATCTCTCTGAGAACATCTGGACAAATGATCAAAAGAAACCAA--- 724

```



QY 202 PheIIePheProAsnThrSerValTyrProGlnuIaIhRgInaGylIeThrThaRgPro 221  
 Db 725 -----GAAAGACAGAAAGATCATATGAAAGAAATTCACACCTACCTCCGTTCT 775  
 QY 222 -----AspLeuPioTyr-GluGlnaI 228  
 Db 776 CATTGAGATTAACAGCAATACATTAGTTTGGCAGACAGAGGGGCCCTATAGAAATGAC 835  
 QY 228 aarGArg-----SerAlaItrPthSer-----HisSerHisProTh 240  
 Db 836 ACACAGAAATTACCCCAAGGCGGCTCTGTGACAGCATGTGCCGCTCCACACCCAGC 895  
 QY 240 rglSerGylalathRgIn-----ProSerE 249  
 Db 896 GTACTCAGCTGTGAGCAGGAAGTTTCAGATGTTCCCAAGTCCGGCTCAGCTCCGTACGC 955  
 QY 249 rSerThr----- 251  
 Db 956 GTCAACSTPACSTGTGTGACGTGACAGACTGCCAGGACAGCACTGAAATTTGTCACCAAC 1015  
 QY 252 -----ValProIysThr-----GluAs -257  
 Db 1016 AATTGTGGAGCTCCCAAGACACAGACTCCCTGAGAACGGTGGGAGACGCTTGAGAGC 1075  
 QY 257 pglAarPProGlnIneAspProTyrGlnIleLeuGlyPro----- 270  
 Db 1076 TCAACACTCCCTCTCCACAGCTCTGGAACGCCAGTCTGCTTGATGTGCAAAGGTT 1135  
 QY 270 ----- 270  
 Db 1136 CCTTCCTTCGAGAGCTTGAGAGATGACTGCAGCCAGTCTCTGCTCCATTAAGCAAC 1195  
 QY 271 -----ThrSerSerArgLeuAlaAsnPro----- 278  
 Db 1196 ATGCTCTTCAAGATATCATCCAGAGAGG-AGTACCCAGCTGGAGCAAGCCAAACAGT 1254  
 QY 279 -----GlySerGlyInIleGlnLeuTPrGlnPh 288  
 Db 1255 TATACSTGCACAGCTGCTGCGCGGCTTCACAGAAATGGAGCTATACGTGTGGCACTT 1314  
 QY 288 eLeuLeuGlnLeuLeuSerAspSerSerAsnSerAsnGylIeThrTrpGlnIyThras 308  
 Db 1315 TCTCTGTGAGCTGATATCAGCAAAATCTCCGACATCATACAGCTGGACTGGAGACGG 1374  
 QY 308 nglGluPheIysMetThrAspProAspRgluValAlaArgArtrPrGluIaGlyIysE 328  
 Db 1375 ATGGGAGTAAAGCTCGCGGCAACCCGATGAGAGTGGCGCGGGGGGAAAGAGGAAAA 1434  
 QY 328 rLysProAsnMetAsnTyrAspLysIleSerArgAlaLeuIaRgTyrTyrTyrAspLysAs 348  
 Db 1435 TAAAGCCCAAGTGAACTAACGAGAAAGCTGAGCGGGGCTTACGCTACTATATACACAMGA 1494  
 QY 348 nIleMetThrLysValIleGlyLysArgTyrAlaTyrLysPhe----- 362  
 Db 1495 CATCATCCACAAAGACGTGGGAGAGCGGTACGTGTACCGCTTCTGTMGCCAGCTCCAGAA 1554  
 QY 363 -----AspPheHisGlyIleAlaGlnIleLeuGlnProHisPr 375  
 Db 1555 CTTCCTGGGGTTCAAGCGCCGAGAACTGACAGCCATCTCTG---GGCGTCCAGCCGACAC 1611  
 QY 375 oPrgIuSerSerMetTyrLysTyrProSerAspLeuPro---ThyMetSerSerTyrH 394  
 Db 1612 GGAGAGACTGAGAGTCCCGGAGACCACTGAGCCCGGCCAGGCTCGTGGACTGAGTGGGA 1671  
 QY 394 salAHisPProGlnIysMet 400  
 Db 1672 AGCCCATCTGACACAGCTG 1690

RESULT 14  
 US-09-920-300A-1693  
 ; Sequence 1693, Application US/09920300A  
 ; Patent No. US20020136728A1

```

: GENERAL INFORMATION:
: APPLICANT: King, Gordon F.
: APPLICANT: Meagher, Madeleine Joy
: APPLICANT: Xu, Jiangchun
: APPLICANT: Secrist, Heather
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.547
: CURRENT APPLICATION NUMBER: US/09/920,300A
: CURRENT FILING DATE: 2001-07-31
: NUMBER OF SEQ ID NOS: 1789
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 1693
: LENGTH: 2268
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-920-300A-1693

Alignment Scores:
Pred. No.:      2,74e-38          Length:      2268
Score:           444.50           Matches:     125
Percent Similarity: 41.28%       Conservative: 43
Best Local Similarity: 30.71%    Mismatches: 105
Query Match:     18.19%         Indels:     135
DB:              11             Gaps:       13

US-09-902-772-2 (1-451) x US-09-920-300A-1693 (1-2268)
QY      122  GUATGATGVALIIEValProLlaAspProThrLeuTrpSerThrAspHisValArgIn 141
Db      571  CAGCGGGCCGTGGGCATTCCAGAAGACCCCCTGGCTGTGGAGTGAGCAGCAAGCATGCCAG 630
QY      142  TTpleuGIUTPrLaIalValLySGluTyrglyLeuProAsPvalAspIleuLeuPheGln 161
Db      631  TGCGTTCCTCGGGCCACCACAAATGATTCAGTCTGTGGGAACGTGAATCTGCAGAGTTCC--- 687
QY      631  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 111
Db      162  AsnIleaspglyLySGluLeuCysLySMetThrLySaSPspPheGlnArgLeuThrPro 181
QY      688  GGCAATGAATGGCCAAATGCTGTGTAACTTGGCAAGAAAGCGCTTCTGGAGCTGGCACCT 747
Db      182  SerTyraSnAlaSpIleuLeuSerHisIleuHISLyLeuArgGluArgGluArgIalThr 201
QY      748  GACTTGTGGGTGATTCCTCTGGGAACAATCTGGAGCAAAATGATCAAGAAAAACCAA--- 804
Db      202  PheIIepHeProASnThrSerValTyrrProGluIalathrGlnArgIllePhthrArgPro 221
QY      805  -----GAAAACACAGACATCAATATGAAAGAAAATTCACACTCCTACCTCCGTTCT 855
Db      222  -----AspleuProTyrr-GluGlnAl 228
QY      856  CATTCGATTAACAGCAATATATTAGTGTTTTGGCAGACAGACAGCGCGCCATTAGGAATGCAG 915
Db      228  aArgArg-----SerIalatrPhrSer-----HisSerHisProth 240
QY      916  ACACAGAAATTAACCCAAAGGCGCCTCTGGACACACATGTGTCGGGCTCCACACACAC 975
Db      240  rGlnSerLySalatrhGln-----ProSerSe 249
QY      976  GTACTCAGACTCTGAGCAGAGAGTTTAGATGTTCCCCCAAGTCTCGGCTCAGCTCCGTACG 1033
Db      249  rSerThr----- 251
QY      1036  GTCACCTACTGCTGTGTCAGTCAGACATTCCACAGCAGCACTTGATTAATTGTCTCAACAC 1099
Db      252  -----ValPolyStrh-----GluAs 257
QY      1096  AATTCTGGAGACTCCCAAGACACACAGCTCCCTTGAGAAAGGTGCGGACAGCTTCGAGAC 1155
Db      257  pGlnArgrProGlnIleuAspProTyrglinIleuEnLyPro----- 270
QY      1156  TCAGACTCCCTCTCCAGTCTGTGGAACACACAGTGTGCTTGTCTGGATGTGCACGGGTT 1211
Db      270  ----- 270
```



Oy	375	oProGlnSerSermethylrlystrprroserAspLeuPro---tyrMetSerSerIyrhi	394
	:::		
Dd	1692	GGAAGACTGAGGTGGCCGGACACCCTGAGGCCGCCCAAGCTCGTGAATGATGGGA	1751
Oy	394	sAlahisrProGlnLysMet	400
		:::::	
Dd	1752	AGCCATCCTGATCCACGCTG	1770

Search completed: July 28, 2003, 10:09:32  
Job time : 297.484 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:30:09 ; Search time 2979.8 Seconds  
(without alignments) 3805.372 Million cell updates/sec

Title: US-09-902-772-2  
2444  
Sequence: 1 MASTIKALSVSEDSLEF.....IYPNTRLPAAHMPSHLGTYY 451

Scoring table: BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=xlh  
-Q=/cgn2\_1/USPto.spool/US09902772/rnps\_23072003\_093658\_14926/app\_query.fasta.1.1294  
-DB=Pending\_Patents\_NA\_Main -QFMT=fastap -SUFFIX=rimpm -MINMATCH=0.1 -IDOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NOM=ext -HEAPSIZE=500 -MINTLEN=0 -MAXLEN=2000000000  
-USER=US09902772.GC.N.1.1.4066.etrnat.23072003.093658.14926 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESUBSTR -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending\_Patents\_NA\_Main.\*

```

1: /cgn2_6/ptodata/1/pna/PCTUS.COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06.COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07.COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08.COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081.COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082.COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083.COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084.COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085.COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086.COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087.COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088.COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089.COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090.COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091.COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092.COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093.COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094.COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095.COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B.COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C.COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US096.COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A.COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B.COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C.COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D.COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E.COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A.COMB.seq.*

```

Result No.	Score	Query Match	Length	ID	Description
1	2444	100.0	1447	US-09-902-772-1	Sequence 1, Appl1
2	2414.5	98.8	1528	US-09-902-772-3	Sequence 3, Appl1
3	2197	89.9	3309	US-09-652-109-9617	Sequence 9617, Ap
4	2197	89.9	3309	US-09-716-953-1936	Sequence 1936, Ap
5	2197	89.9	3309	US-09-721-589-5458	Sequence 5458, Ap
6	2197	89.9	29	US-09-726-806-4625	Sequence 4625, Ap

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

```

7 2197 89.9 3309 29 US-09-726-811-4920 Sequence 4920, Ap
8 2182 89.3 1389 35 US-09-948-941-233 Sequence 233, App
9 2182 89.3 3166 1 PCT-US02-23913-98 Sequence 98, Appl
10 2182 89.3 3166 30 US-09-784-356-26 Sequence 26, Appl
11 2182 89.3 3166 32 US-10-021-660-26 Sequence 26, Appl
12 2182 89.3 3166 42 US-10-205-823-98 Sequence 98, Appl
13 1991 81.5 2172 80 US-09-360-207-355 Sequence 355, App
14 1810 74.1 3126 18 US-09-442-384A-475 Sequence 475, App
15 1810 74.1 3126 18 US-09-442-384B-475 Sequence 475, App
16 1624.5 66.5 1359 38 US-10-035-832-1210 Sequence 1210, App
17 1624.5 66.5 1359 39 US-10-052-482-198 Sequence 198, App
18 1624.5 66.5 2938 7 US-08-343-443-9 Sequence 3, Appl1
19 1624.5 66.5 2957 1 PCT-US02-18947-710 Sequence 710, App
20 1624.5 66.5 2957 38 US-10-007-926A-295 Sequence 295, App
21 1624.5 66.5 2957 38 US-10-035-832-1209 Sequence 1209, App
22 1624.5 66.5 2957 39 US-10-052-482-197 Sequence 197, App
23 1624.5 66.5 2957 41 US-10-172-118-710 Sequence 710, App
24 1624.5 66.5 3298 25 US-09-652-814-9467 Sequence 9467, App
25 1624.5 66.5 3298 27 US-09-699-999-6013 Sequence 6013, App
26 1624.5 66.5 3298 28 US-09-716-953-1954 Sequence 1954, App
27 1624.5 66.5 3298 28 US-09-716-990-1139 Sequence 1139, App
28 1624.5 66.5 3298 29 US-09-721-589-5554 Sequence 5554, App
29 1624.5 66.5 3298 29 US-09-726-171-2148 Sequence 2148, App
30 1624.5 66.5 3298 29 US-09-726-806-4681 Sequence 4681, App
31 1624.5 66.5 3372 17 US-09-396-970-7615 Sequence 7615, App
32 1613.5 66.0 1359 38 US-10-035-832-1207 Sequence 1207, App
33 1613.5 66.0 1359 38 US-10-052-482-195 Sequence 195, App
34 1613.5 66.0 1729 14 US-09-053-832-1206 Sequence 627, App
35 1613.5 66.0 1729 38 US-10-035-832-1206 Sequence 1206, App
36 1613.5 66.0 1729 39 US-10-052-482-194 Sequence 194, App
37 1613.5 66.0 1729 80 US-09-360-207-10845 Sequence 10845, A
38 1611.5 65.9 3110 76 US-06-324-185-33223 Sequence 33223, A
39 1595 65.3 3107 61 US-06-172-373-5267 Sequence 5267, App
40 1536.5 62.9 1932 14 PCT-US02-25766-2596 Sequence 2596, App
41 1536.5 62.9 1932 14 US-09-053-3758-327 Sequence 327, App
42 1536.5 62.9 1932 18 US-09-442-384A-530 Sequence 530, App
43 1536.5 62.9 1932 18 US-09-442-384B-530 Sequence 530, App
44 999 40.9 661 29 US-09-726-806-3434 Sequence 3434, App
45 965 39.5 1431 1 PCT-US02-06518-147 Sequence 147, App

```

## ALIGNMENTS

```

RESULT 1
US-09-902-772-1
: Sequence 1, Application US/09902772
: GENERAL INFORMATION:
: APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
: FILE REFERENCE: Chugai selyaku kabushiki kaisha 5001
: CURRENT APPLICATION NUMBER: US/09/902,772
: PCT FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: US/08/878,177
: PRIOR FILING DATE: 1997-06-18
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1447
: TYPE: DNA
: ORGANISM: C-11 gene, c-erg gene w/ deletion, chicken DNA
US-09-902-772-1

```

## Alignment Scores:

```

Pred. No.: 2.06e-181 Length: 1447
Score: 2444.00 Matches: 451
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 34 Gaps: 0

```

US-09-902-772-2 (1-451) x US-09-902-772-1 (1-1447)

```

Qy 1 MetalaserThrilleygluAlaLeuSerValSerGluInsproGlnSerLeuPheGlu 20
Db 63 ATGCACAGCACTATTAAAGCAATATATGATGATGATGAAGCAAGCTCTGTTGAG 122
Qy 21 CysAlaIyrglySerProHlsLeuAlaIySthrGluMetThrAlaSerSerSerGlu 40
Db 123 TGTCCACAGATGATGCCCACTTGCACAAACAGAAATGACAGCTCTTCCAGTAA 182
Qy 41 TyrcIleThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnPro 60
Db 183 TATGGCAACATATAAAGATAGAGCCCGGCTTCCACAGAGAGCTGTTATCAGACC 242
Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db 243 CCGGCACAGATTACATTAAGATGAGATGATTAACCAACAGGTTAATGATGATGAAGAA 302
Qy 81 SerProAspAspCysSerValAlaIyGlyGlyMetValSerSerSerAspVal 100
Db 303 TCACCTGATGATCCTCAGCGTGCACAAAGAGGAAATGTTACACTTACAGACAAATGTT 362
Qy 101 GlyMetAsnTyrglySerTyrmGluGluLysHisIleProProProAsnMetThr 120
Db 363 GGGATGACATCTGAGACATGATGAGAGAGAGATATTCGCTCCAAATATGACAAAC 422
Qy 121 AsnGluArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
Db 423 AATACAGAAAGATTATGTCGACAGATCTCTGTTATGAGACACAGACATGTCAGG 482
Qy 141 GlnThrLeuGluThrAlaValAlaIyGlyGlyLeuProAspValAlaIleLeuPhe 160
Db 483 CAGTGGGTGGTGGGAGTGAAGAGATATGCTTCCAGCGTGCATCTGTTGTTTC 542
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 543 CAGACACTTGTATGGAGAAAGCTGTGTAATGACCAAAATGACTTCCAGAGACTACG 602
Qy 181 ProSerTyrsAlaAspIleLeuLeuSerHisLeuHisIyThrLeuArgIyGlyAla 200
Db 603 CCGAGCTATTAAGCAGATATCTCTGTCACACTTACATCACTCAGAGAGAGAGGCC 662
Qy 201 ThrPheIlePheProAsnThrSerValTyrrProGluAlaThrGlnArgIleThrArg 220
Db 663 ACTTTATTTTCCAAATATACATCATGTTTACCACCAACAGCAAAATAAACAAAG 722
Qy 221 ProAspLeuProTyrgluGlnAlaIyArgSerAlaThrPheSerHisSerHisProThr 240
Db 723 CCAATTTACTTATGAGCAAGCAGAGATGACCGTGCAGAGATCAGACATCCACT 782
Qy 241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGlnAspGlnArgPro 260
Db 783 CAGTCAAAAGCTACCAACCATCATCTTCAACAGTGCACCAAAACAGAAACAGCCGTC 842
Qy 261 GlnLeuAspProTyrglnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
Db 843 CAGTTAATCTTATTCAGATCTTGTGACCAACAGCCGCTTGTGAAATCCAGGGAGT 902
Qy 281 GlyGlnIleGlnLeuThrPglInPheLeuLeuGluLeuLeuSerAspSerSerAsn 300
Db 903 GGGCAGATACAGCATATGAGCATCTCTACTGAGAGCTTGTGCGACACTCCAACTCAAC 962
Qy 301 CysIleThrTrpGluGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAla 320
Db 963 TGCATCACCTGGGAGGACCAAAATGGGAGTTCAAAGATGACAGCCCTGTAAGTGGCT 1022
Qy 321 ArgArgTrpGlyGluAlaGlySerLysProAsnMetLysAsnTyrrAspLysLeuSerArgAla 340
Db 1023 CCGCGTGGGAGAGAGAAACCAACCTAACATGAATATGACAAACCTACACCGGCA 1082
Qy 341 LeuArgTrpTyrrTyrrAspLysAsnIleMetThrLysValHisGlyLysArgTyrrAlaTyrr 360
Db 1083 CTTCGCTACTACTATGACAAATATTTATGACTTAAGTTATGATTAACCTATGCTTAC 1142
Qy 361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerSerMet 380

```

```

|||||
Db 1143 AATATTGATTTCCAGGAAATGCGTCAGGCCCTCCAGGCTCCAGCTCCAGATCATG 1202
Qy 381 TyrlstfyrProserAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMet 400
Db 1203 TACAAATACCATCATGAGCTCCCTCCATGAGTTCATCCATGACACCCCAAGATG 1262
Qy 401 AspHeValAlaProHisProAlaLeuProValThrSerSerPhePheAla 420
Db 1263 AACTTTGAGTCCCATCCCTCCCTGCTTGGCCCTAACCTCATGACGCTTTTGGCTGCC 1322
Qy 421 ProAsnProTyrTTPasSerProThrGlyLysLeuTyrProAsnThrArgLeuProAla 440
Db 1323 CCTAATCCATCATCTGGAATTCACCAATGAGAGCATCTACCCCAATACAGGCTCCAGCT 1382
Qy 441 AlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1383 GCTCATATGCTCTCCATCTTGCGACCTACTAC 1415

```

## RESULT 2

```

US-09-902-772-3
: Sequence 3, Application US/09902772
: GENERAL INFORMATION:
: APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
: FILE OF INVENTION: the Proteins
: FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001
: CURRENT APPLICATION NUMBER: US/09/902,772
: CURRENT FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: US/08/878,177
: PRIOR FILING DATE: 1997-06-18
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 1528
: TYPE: DNA
: ORGANISM: c-erg gene, chicken DNA
US-09-902-772-3

```

## Alignment Scores:

```

Pred. No.: 4.53e-179 Length: 1528
Score: 2414.50 Matches: 450
Percent Similarity: 94.14% Conservative: 0
Best Local Similarity: 94.14% Mismatches: 1
Query Match: 98.79% Indels: 27
DB: 34 Gaps: 1

```

US-09-902-772-2 (1-451) x US-09-902-772-3 (1-1528)

```

Qy 1 MetAspSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 63 ATGCCAAGCACTATTAAAGAAAGCACTTATCAGTGGTGAAGTGAAGCAACGCTTTTGGAG 122
Qy 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGlu 40
Db 123 TGTGCTACGAGATGCGCCCACTTGCAGAAAGCAAGAAATGACAGCTCTCTCCAGTGA 182
Qy 41 TyrGlnInthrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 60
Db 183 TATGGGCAACATCAAGATGAGCCCGGCTTCCCGACAGCAGCTGTTATCAACAGCCC 242
Qy 61 ProAlaThrValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db 243 CCGGCCAGATTCACATTAAAGATGAGTGAAGTGAAGCAACAGGTTAATGGGTCAAGGAAT 302
Qy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db 303 TCACCTGATGCTGACACGCTGGCAAAAGGAGGAAATGTTAGCAATTCAGCAACATGTT 362
Qy 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThr 120
Db 363 GGGATGAACATATGGAACCTACATGGAAGAGAGATATTCGCCCTCAATATGACAAACC 422

```

```

Qy 121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrPheSerThrAspHisValArg 140
Db 423 AATGACGAAGAGATTATTGTCGCCAGATCTTACGATTATGAGACACAGACATGTACGG 482
Qy 141 GlnTyrPleuGlnTyrPalValLysGluTyrGlyLeuProAspValAspIleLeuLysPhe 160
Db 483 CAGTGGCTGGAGTGGCAGTGAAGAGATATGCTTCCAGACGTCGACATCTTGTGTTCC 542
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 543 CAGAACATTTGATGGGAAAGATTGTGTAATGACCAAAAGAGACTTCCACAGACTCAGC 602
Qy 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg ----- 198.
Db 603 CCGAGCTATACGAGATATCTCTGTCACACCTACACTTCCACAGAGACATCCCTT 662
Qy 198 ----- 198
Db 663 CCACATTTGACTTCAGATGATGTTGATTAAGGCTTACAAAACCTCCACGGTTAATGCAT 722
Qy 199 -----GlyAlaThrPheLeuPheProAsnThrSerValTyrProGluAla 213
Db 723 GCTAGAAACACAGGAGAGCCACTTTATTTTCCAAATATCATAGTTTACCCAGAAACA 782
Qy 214 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyr 233
Db 783 ACGAAAGAAATACACACAGGCCAGATTTCCTTATGACAGAGAGAGATCAGCTGG 842
Qy 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 253
Db 843 ACGAGTCACAGCCATCCCACTCAGTCAAAAGCTACCAACATCATCTTCAACAGTCCCC 902
Qy 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
Db 903 AAACACAGAACACAGCTCTCATCTTATGATTCCTTATGATTCCTTGAACGACGACGAGC 962
Qy 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeu 293
Db 963 GCTTCTCAAAATCCAGGAGATGGGCAATACAGTATGCGAGTTCCTACTGAGACTCTG 1022
Qy 294 SerAspSerSerAsnSerAsnCysIleThrTyrGlnGlyThrAsnGlyLysPheLysMet 313
Db 1023 TCGGACAGCTCCCAACTCAACTGATCATCTGGAGGAGGACAAATGGGAGTTCAAGATG 1082
Qy 314 ThrAspProAspGluValAlaArgTyrGlyLysLysSerLysProAsnMetAsn 333
Db 1083 ACAGACCTGATGAAGTGGCTCGGCTGGGAGAGAGGAAACCAACCTAACATGTAAC 1142
Qy 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353
Db 1143 TATGACAAACTCAGCCGCTGCACTTGCCTACTACTATGACAAAATATTATGACTAAAGTT 1202
Qy 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
Db 1203 CATGCTAAACGCTATGCTCAAAATTTGATTTCCACGAAATGCGTCAAGCCCTCCAGCT 1262
Qy 374 HisProProGlnLysSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
Db 1263 CACCCCTCAAAATCATCATGATCAATTAACCATCAACACCTCCCTCATGAGATTCCTAC 1322
Qy 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
Db 1323 CATGACACCCCAAGAGATGAACCTTGTAGCTTCCCATCCCTGCTTGGCCCGTAACC 1382
Qy 414 SerSerSerPhePheAlaAlaProAsnProTyrTTPasSerProThrArgIlyLysTyr 433
Db 1383 TCATCCAGCTTTTGTGTCCTTAATCCATACAGAAATTCACCAACTGAGAGGATCTAC 1442
Qy 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1443 CCCAATACAGGCTGCGACGCTGCTCATATGCTTCCATCTTGGCACCTACTAC 1496

```

## RESULT 3

US-09-652-109-9617  
 ; Sequence 9617, Application US/09652109  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McCarthy, Sean A.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
 ; TITLE OF INVENTION: THEREFOR  
 ; FILE REFERENCE: 1600.1180-001  
 ; CURRENT APPLICATION NUMBER: US/09/652.109  
 ; CURRENT FILING DATE: 2000-08-30  
 ; PRIOR APPLICATION NUMBER: 60/151,128  
 ; PRIOR FILING DATE: 1999-08-30  
 ; NUMBER OF SEQ ID NOS: 10105  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9617  
 ; LENGTH: 3309  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-652-109-9617

## Alignment Scores:

Pred. No.:	1,31e-161	Length:	3309
Score:	2197.00	Matches:	410
Percent Similarity:	91.96%	Conservative:	13
Best Local Similarity:	89.13%	Mismatches:	23
Query Match:	89.89%	Indels:	14
DB:	25	Gaps:	4

US-09-902-772-2 (1-451) x US-09-652-109-9617 (1-3309)

QY 1 MetLaserhrilleysgluaaleuSerValuaSerGluaspGlnserleupheglu 20  
 DB 128 ATGGCCGACACTATTAAAGAGCCTTATCAGTTGTGTGAGGAGCCAGTCTTTTGTAG 187  
 QY 21 CysAlaTyrglySerProHisleuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 DB 188 TGTGCTACAGGAAGCAGACACCTGTGCTAAGCAGAGATGACCGGCTCTCTCCAGGAC 247  
 QY 41 TyrglylnthrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 DB 248 TATGACAGACTTCCAGATGAGCCACGCGCTCCCTGAGCAGATGCTGTCTCAACCC 307  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProHisGlnValAsnGlySerArgAsn 80  
 DB 308 CCACCCAGGGCTCACCATCAAAATGATGATACCCCTAGCCAGGTAATGCTCAAGGAGAC 367  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 DB 368 TCTCTGATGATGACATGCTGTGGCCAAAGGGGGAAGATGTTGGGACCCAGACCGCTT 427  
 QY 101 GlyMetAsnTyrglySerTyrrMetGluGluLysHisIleProProProAsnMetThr 120  
 DB 428 GGGATGACACTACGCGACCTACATGAGAGAGACACATGCGTCAACCCCAACATACACAGC 487  
 QY 121 AsnGlnArgValIleValAlaProAlaAspProThrLeuTrpSerThrAspHisValArg 140  
 DB 488 AAGCAGCGCAGAGATTATCTGCGCCAGCATCTCTAGTATGAGATACAGACCATGTGCGG 547  
 QY 141 GlnTrpLeuGlnTrpAlaValLysGlyLysLeuProAspValAlaIleLeuLeuPhe 160  
 DB 548 CAGTGGCTGAGTGGGGGTGAAGATATAGCCTTCCACAGCCTCACTTGTATTC 607  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspArgGlnArgLeuThr 180  
 DB 608 CAGAACATCGATGGGAGGAACTGTGCAAGATGACCAAGAGCGACTTCCAGAGGCTCAC 667  
 QY 181 ProSerTyrrAsnAlaAspIleLeuLeuSerHisIleuHisTyrrLeuArgGlnArgGlyAla 200  
 DB 668 CCCAGCTACAAAGCCGACATCTTCTCTACATCTCCACTACCTCAAGAGACCTCTT 727  
 QY 201 ThrPheIlePheProAsnThrSerValTyrrProGlnAlaThrGlnArgIleThrThrArg 220  
 DB 728 CCACATTTG-----ACTTCA-----GATCATGTGTGATTAAGCCTTACAAACTCT 772

QY 221 Pro-----AspLeuProTyrgluGlnAlaArgSerAla 232  
 DB 773 CCACGGTAAATGATGCTAGAAACACAGATTATACATATGACCCCGCCAGAGATCAGCC 832  
 QY 233 TrpThrSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThr 251  
 DB 833 TGGACCGGTACAGCGCCACCCACGACCCCACTCGAAGCTCTCAACATCTCTCTCCACA 892  
 QY 252 ValProLysThrGlnAspGlnArgProGlnLeuAspProTyrglnIleLeuGlyProThr 271  
 DB 893 GTGCCCAAACTGAAGACAGCGCTCCACTATGATCTTATCAGATCTTGGACACACA 952  
 QY 272 SerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGln 291  
 DB 953 AGTAGCGCGCTTGAATCCAGGACAGTGGCCAGATCCAGCTTGGCGAGTTCTCTCGAG 1012  
 QY 292 LeuLeuSerAspSerSerAsnSerAsnGlyIleThrTrpGlnGluTrpAsnGlyLysPhe 311  
 DB 1013 CTCTGTGAGACACTTCAACTCCAGCTGATACCTGGGAGAGCACCACCAAGGGAGTTTC 1072  
 QY 312 LysMetThrAspProAspGlnValAlaArgArgTrpGlyLysLysSerLysProAsn 331  
 DB 1073 AAGATGACGATCCCGACAGAGTGGCCCGGCGTGGGAGAGCGGAAGACCAACCCAAC 1132  
 QY 332 MetAsnTyrrAspLysLeuSerArgAlaLeuArgTyrrTyrrAspLysAsnIleMetThr 351  
 DB 1133 ATGAACCTACATATAGCTACACCGCGCCCTCCGTACTACTATGACCAAGAACATCATACC 1192  
 QY 352 LysValHisGlyLysArgTyrrAlaTyrrLysPheAspPheHisGlyIleAlaGlnAlaLeu 371  
 DB 1193 AAGGTCCATGGGAAGCGCTACCGCTTACAGTTGCACTTCCACGGGATCGGCCAGCCCTC 1252  
 QY 372 GlnProHisProGlnLysSerMetTyrrLysTyrrProSerAspLeuProTyrrMetSer 391  
 DB 1253 CAGCCACACCCCGGAGTATCTGTACAAAGTACCCCTCAGACCTCCCGTACATGGGC 1312  
 QY 392 SerTyrrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuPro 411  
 DB 1313 TCTTATACCCGCCACCAAGAGATGAACTTGTGTGGCCGCCACCTCCAGCCCTCC 1372  
 QY 412 ValThrSerSerPhePheAlaAlaProAsnProTyrrTrpAsnSerProThrGlyGly 431  
 DB 1373 GTGACATCTTCCAGTTTCTTCTGCTCCCAACCCATCTGATTCACCACTGGGGGT 1432  
 QY 432 IleTyrrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrr 451  
 DB 1433 ATATACCCCAACTAGCTTCCACACGACGATATGCTCATCTGTGGGCACTTACTAC 1492

RESULT 4  
 US-09-716-953-1936  
 ; Sequence 1936, Application US/09716953  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gearing, David P.  
 ; APPLICANT: Holtzman, Douglas A.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
 ; TITLE OF INVENTION: THEREFOR  
 ; FILE REFERENCE: 1600.2044-001  
 ; CURRENT APPLICATION NUMBER: US/09/716.953  
 ; CURRENT FILING DATE: 2000-11-21  
 ; PRIOR APPLICATION NUMBER: 60/167,413  
 ; PRIOR FILING DATE: 1999-11-24  
 ; NUMBER OF SEQ ID NOS: 2620  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1936  
 ; LENGTH: 3309  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-716-953-1936

Alignment Scores:  
 Pred. No.: 1,31e-161 Length: 3309  
 Score: 2197.00 Matches: 410



QY		1	MetalSerThrIleValSerGluLeuSerValValSerGluAspGlnSerLeuPheGlu	20
Db		128	ATGGCAGACACTATTAAAGAACCTTATCACTGTTGATGTAGAACAACGTCCTTGTTGAG	187
QY		21	CysAlaTyrglySerProHisLeuAlaIsthrgLueThraIasSerSerSerglu	40
Db		188	TGTGCCTTAAGGAAGCGCACACTGGCTTAAGACAGAGATGACGGGCTCTCTCCAGCGAC	247
QY		41	TyrGIgIInThSerIstMetSerProArGvaLPproGlnIInAsPTPLenSergInPro	60
Db		248	TATGGACAGACTCCCAAGATGAGGCCACGGCTCCCTCAGCAGATTGGCTTCTCAAACC	307
QY		61	ProAlaArgValThrIleIstMetGluCysAsnProAsnGlnValAsnglySerArgSn	80
Db		308	CCAGCCAGGCTACCATCAAAATGGAATGTAACCTTACCGACGGTAATGGCTCAAGAAC	367
QY		81	SerProAspAspCysSerValAlaIstGlyIstMetValSerSerSerAspAsnVal	100
Db		368	TTCTCTGATGAAGACGTGTGGCCAAAGGCGGGAAGATGGGGCAGCCAGACACGGT	427
QY		101	GlyMetAsnTyrglySerIstMetIugIulustHisIleProproProAsnMetThcTr	120
Db		428	GGGATGAATACGGGAGCTACATGTGAGGAGAAAGACATGCCACCCCAACATATGACAG	487
QY		121	AsnGluArgArgValIleValProAlaAspProthrIeUtpSerThrAspHisValArg	140
Db		488	AACGACCGAGAGATTATCGTCGACACAAATCCTTAGCTATGAGATAAGACAACTGGCG	547
QY		141	GlnTripleGltTriplaValIstGlyIulArgIleUpRoAspValAspIleLeuIlePhe	160
Db		548	CATGGCTGTAGTGGGCGGTGAANAARITGGCTTCAGACGTCAACATCTTGTATTC	607
QY		161	GlnAsnIleAspGlyIstGluIstMetThryIstAspAspPheGlnArgLeuThr	180
Db		608	CAGAACATCGATGGAGAACTGTGCAAGATGACCAAGACAGACTTCCAGAGGCTCAC	667
QY		181	ProSerTyrAsnAlaAspIleLeuIstMetHISLeuHISTryrLeuArgGluArgGlyAla	200
Db		668	CCGAGTAAACGGCGACATCCTTCTCAACATCCACATCCACAGAGAACTCCTT	727
QY		201	ThrPheIstPheProAsnThrservalTyrrProGluAlaThrGlnArgIleThrThrArg	220
Db		728	CCACATTG-----ACTTCA-----GATGATGTGATTAAGCCTTACAAAACCT	772
QY		221	Pro-----AspLeuProtyrGluGlnAlaIstArgSera	232
Db		773	CCAGCGTTATGATGATCTGAACAACAAATTACATATGAGCCCCAGAGAGATCAGCC	832
QY		233	TriPThrSerHisSerIstProthr---GlnserIstAlaThrGlnProSerSerSerThr	251
Db		833	TGACCGGTGACGGGACCCACGCCCAAGTCGAAGTGTCTCAACATCTCTCTCCACA	892
QY		252	ValIstProIstThrgIuAspGlnArgProGlnIstMetAspProtyrGlnIleLeuGlyProThr	271
Db		893	GTCGCCAANAACGTAGAGACAGCGTCTCAGTTAGATCTTATCGAATTCCTTGACCAACA	952
QY		272	SerSerArgIleuAlaAsnProIstGlyGlnIleGlnIstMetArgPheIleuLeuGln	291
Db		953	AGTAGCGGCTTGCAATCAACAGCAGTGGCCAGATCCAGCTTGGCAGTTCCTCTGGAG	1011
QY		292	LeuIstSerAspSerSerAsnSerAncysIleThrTriPGlugIyThrAsnGlyGluPhe	311
Db		1013	CTCCTGTGCGAAGCTCCAATCTCAGCTGACATCACTGTGGAAGGCAACCAAGGGGATTC	107
QY		312	LysMetThrAspProAspGluValAlaIstArgArgTriPGlyGluArgIstMetLysProAsn	331

## RESULT 5

; GENERAL INFORMATION:

APPLICANT: Villeva, Jean-Lu

1111 OF INVENTION. THEREON  
 1111 1500 2045 00

FILE REFERENCE: 1600.2045-001

CURRENT APPLICATION NUMBER: 05/09/121,585

PRIOR APPLICATION NUMBER: 60/116

PRIOR FILING DATE: 1999-11-24

NUMBER OF SEO ID NOS: 7017

```

; SOFTWARE: FastSEQ for Windows Version 4.0

```

; SEQ ID NO 5

; LENGTH: 3309

TYPE: DNA

ORGANISM: Homo sapiens

05-09-721-589-5498

**Aljama + Ecorra:**

Alignment scores:  
Pred NO : 1

Score: 21

Percent Similarity: 91

Best Local Similarity: 89

Query Match: 89

DB: 29

PC-00-003-773-3 / 1-151) x

X (TC4-T) 7-711-706-60-60

1 MetAlaSerThr  
OY

$r \times$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Db 128 ATGCCCAGCACT

QY 21 CysAlaTyrGly

Dh  
100

DD 100 1616111AUG 68

41 TVRGIVG1nThr

	$\gamma_{\infty}$
$\vdots$	$\vdots$
$\gamma_1$	$\gamma_0 = \gamma_{\infty} - 1$

Db 248 TATGGACAGACT

QY 61 ProLaArgVal

309

DD 308 CCAGCCAGGTC





Db 1013 CTCCTGTGCGAGAGCTCCCACTCCAGCTGCATACCTTGGAGAGCCACAGGGAGTTTC 1072  
Qy 312 LysMetThrAspProAspGluValAlaIArgArgTrpGlyGluArgLysSerLysProAsn 331  
Db 1073 AAGATGACGATGCCGAGAGAGTGGCCGGCTGGGAGAGCGGAAGACCAACCCCAAC 1132  
Qy 332 MetAsnTrpAspLysLeuSerArgAlaLeuArgTrpTrpTrpAspLysAsnLleMetThr 351  
Db 1133 ATGAACCTAGATGAGCTAGCGGGCCCTCCGTTACTACTATGACAAACATCATGTAGCC 1192  
Qy 352 LysValHisGlyLysArgTrpAlaTrpLysPheAspPheHisGlyLleAlaGlnAlaLeu 371  
Db 1193 AAGGCTCATGGAGAGCGCTAGCGCTACAGAGTTCCAGCTCCAGGGATGCCAGGCCCTC 1252  
Qy 372 GlnProHisProProGluLysSerSerMetLysTrpProSerAspLeuProTyrMetSer 391  
Db 1253 CAGCCCCCAGCCCCGAGATCATCTCTGTACAGATGACCCCTCAGACCTCCCTACATGGCC 1312  
Qy 392 SerTrpHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuPro 411  
Db 1313 TCCTATACGCCACCCACAGAGATGACTTTGGGGCCCCCAGCCCTCCAGCCCTCCCC 1372  
Qy 412 ValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGly 431  
Db 1373 GTGACATCTTCAGTTTTCCTGCTGCCCCCAACCATCTGGAATTCACCAACTGGGGGT 1432  
Qy 432 IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTrp 451  
Db 1433 ATATACCCCAACACTAGGCTCCCGACGACATATGCTCTCATCTGGGACATTACTAC 1492

## RESULT 8

US-09-948-941-233  
; Sequence 233, Application US/09948941  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO00788  
; CURRENT APPLICATION NUMBER: US/09/948, 941  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/231,328  
; NUMBER OF SEQ ID NOS: 12618  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 233  
; LENGTH: 1389  
; TYPE: DNA  
; ORGANISM: Human  
US-09-948-941-233

## Alignment Scores:

Pred. No.: 6,89e-161 Length: 1389  
Score: 2182.00 Matches: 407  
Percent Similarity: 91.72% Conservative: 14  
Best Local Similarity: 88.67% Mismatches: 24  
Query Match: 89.28% Indels: 14  
Gaps: 4

US-09-902-772-2 (1-451) x US-09-948-941-233 (1-1389)

Qy 2 AlAserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLysPheGluLys 21  
Db 25 GCACCTCATATCAAGGAAGCCCTTATCACTGTGTGAGTGAGGACCACTGCTTTTGTAGTGT 84  
Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41  
Db 85 GCCACGGAAGACCAACCTTGCTTAAGACAGATGACCGCGTCTCTCTCCAGGACATAT 144  
Qy 42 GlycineThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61  
Db 145 GGAACAGACTTCCAGAGTAGAGCCACGCGTCCCTCAGCAGATGCGTGTCTCAACCCCA 204  
Qy 62 AlaArgValThrIleLysMetGluLysAsnProAsnGlnValAsnGlySerArgAsnSer 81

Db 205 GCCAGGGTCACCATCAAAATGGAATGTAACTTACCCCTAGCCAGGAGTAAAGTCTCAAGAACTCT 264  
Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101  
Db 265 CCTGATGATGACAGTGTGGCCAAAGCCGGAAGATGTGGGGACACCCAGACACCGTTGGG 324  
Qy 102 MetAsnTrpGlySerTrpMetGluGluLysHisIleProProProAsnMetThrAsn 121  
Db 325 ATGAACCTAGGACCTACATGAGAGACACATGTGCCACCCCAACATCACCACAGAAC 384  
Qy 122 GluArgArgValIleValProAlaAspProThrIleLeuTrpSerTrpAspHisValArgGln 141  
Db 385 GAGGCGAGAGTTATGCTGCGCAGCAGATCTCAGCTGAGATGAGATGAGACATGTCGGGCG 444  
Qy 142 TrpLeuGluTrpAlaValLysGlyLysGlyLysLeuProAspValAspIleLeuLeuPheGln 161  
Db 445 TGGCTGAGAGTGGGGGTGAAGAAATATGCGCTTCCAGACGTCACATCTGTATTTCAG 504  
Qy 162 AsnIleAspGlyLysGluLeuLysGlyMetThrLysAspAspPheGlnArgLeuThrPro 181  
Db 505 AACATCGATGGGAAGAACTGTGCAGATGACCAAGGACGACTTCCAGAGGCTCACCCCC 564  
Qy 182 SerTrpAsnAlaAspIleLeuSerHisIleuHisTrpLeuArgGluArgGlyAlaThr 201  
Db 565 AGCTACAAAGCCGAGACATCTTCTCATATCTCCAGTCACTCAGAGAGACTCTCTTCCA 624  
Qy 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgPro 221  
Db 625 CATTTG-----ACTTCA-----GATGATGTGATTAAGCCTTACAAAACTCTCCA 669  
Qy 222 -----AspLeuProTyrGluGlnAlaIArgArgSerAlaTrp 233  
Db 670 CGGTTAATGCTAGTACAAACACAGATTTTACATATGAGCCCCACAGAGATGACCTGG 729  
Qy 234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 252  
Db 730 ACCGGTACAGGCCACCCACGCGCCAGCGAGCGAAGCTGCTCAACCATCTCTCCACAGTG 789  
Qy 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272  
Db 790 CCCAAACTGAAAGACAGCGCTCAGATTAATCTTATGATGATTTGGACCAACAGT 849  
Qy 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeu 292  
Db 850 AGCCGCTTGCAGAAATCCAGGACGAGCGCAGATCCAGCTTGGCAGTTCTCTCGAGCTC 909  
Qy 293 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGluThrAsnGlyLysPheLys 312  
Db 910 CTGTGGACAGCTCACTCACTGCTGATCAGCTGGGAGAGACCAACGGGAGTTCAAG 969  
Qy 313 MetThrAspProAspGluValAlaIArgArgTrpGlyGluArgLysSerLysProAsnMet 332  
Db 970 ATGACGAGATCCCGACGAGAGTGGCCCGCGCTGGGAGAGCGGAAGACCAACCAACATG 1029  
Qy 333 AsnTrpAspLysLeuSerArgAlaLeuArgTrpTrpTrpAspLysAsnLleMetThrLys 352  
Db 1030 AACTACGATTAAGCTCACCGCGCCCTCGTTACTACTATGACAAACATCATCATGACCAAG 1089  
Qy 353 ValHisGlyLysArgTrpAlaTrpLysPheAspPheHisGlyLleAlaGlnAlaLeuGln 372  
Db 1090 GTCCATGGGAAGCCCTAGCGCTTACAGATGTCAGACTTCCACGGGATGCCAGGCCCTCAG 1149  
Qy 373 ProHisProProGluLysSerMetLysTrpLysTrpProSerAspLeuProTyrMetSerSer 392  
Db 1150 CCCACACCCCGAGGATATCTCTGTAAAGTAAACCCCTCAGACCTCCGCTATAGGGCTCC 1209  
Qy 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412  
Db 1210 TATCAGCCCAACCAACGAGAGATGAATCTTGTGGCGCCCAACCTCCAGGCTCCCGGTG 1269  
Qy 413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIle 432

```

Db      1270 ACATCTTCAGTGTTCCTGCTGCCCAAAACCCATGAAATTCACCACTGGGGGATA 1329
QY      433 TTPProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db      1330 TACCCCAACACTAGGCTCCCAACAGCCATATGCTTCATCTGGGCACTTACTAC 1386

RESULT 9
PCT-US02-23913-98
; Sequence 98, Application PC/TUS0223913
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-04APC
; CURRENT APPLICATION NUMBER: PCT/US02/23913
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-23913-98

Alignment Scores:
Pred. No.:      1,86e-160      Length:      3166
Score:          2182.00      Matches:      407
Percent Similarity: 91.72%      Conservative: 14
Best Local Similarity: 88.67%      Mismatches:  24
Query Match:    89.28%      Indels:      14
                        Gaps:      4

US-09-902-772-2 (1-451) x PCT-US02-23913-98 (1-3166)

QY      2 AlaSerThrIleLysGluAlaLeuSerValSerGluAspGlnSerLeuPheGlyCys 21
Db      281 GCAGCTCATATCAAGAGAGAGCCCTTATCATGTTGAGAGAGACAGTCGTTGTTGAGTGT 340
QY      22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlyIle 41
Db      341 GCCTTAGGGAAGCGCACACCTGGCTTAAGACAGAGATGACCGCTCTCTCCAGCGACTAT 400
QY      42 GlyIleThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 61
Db      401 GGACAGACTTCCAAATGAGCCACCGCGTCCCTCACAGAGATTGGCTGTCTCAACCCCA 460
QY      62 AlaArgValThrIleLysMetGlyCysAsnProAsnGlnValAsnGlySerArgAsn 81
Db      461 GCCAGGGTCAACATCAAAATGGAATGTAACCCCTAGCCAGGTGAATGCTCAAGAGACTCT 520
QY      82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnValGly 101
Db      521 CCGATGATGATGCGAGTGTGGCCAAAGCGGGAGAGATGTGTGGGCGAGCCAGACCGTTGG 580
QY      102 MetAsnTyrGlySerTyrMetGluGlyLysHisIleProProProAsnMetThrTrpAsn 121
Db      581 ATGAAGTACGGCGAGTACATGAGGAGAGACATCCACCCCAACATGACACGAGAAC 640
QY      122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db      641 GACCGCAGAGTATGCTGCCAGCAGAGATCTACAGCTATGAGATACAGACCATGTGCGCAG 700

QY      142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161
Db      701 TGGCTGAGTGGGCGGTGAAGATATGCTTCCAGACTCATCTTGTATTCCAG 760
QY      162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db      761 AACATGATGGAGGAAGACTGTGCAGATGACCAAGAGACACTTCCAGAGACTCACCC 820
QY      182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
Db      821 AGCTAACAGCCGACATCTCTCTCAATCTCCATCTCCAGAGACTCTCTTCCA 880
QY      202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrTrpArgPro 221
Db      881 CATTTG-----ACTTCA-----GATGATGTGATTAACCTTACAAACTCTCCA 925
QY      222 -----ASpLeuProTyrGlnGlnAlaArgArgSerAlaTrp 233
Db      926 CGGTTAATGATGCTGTAAGAAACACAGATTTACCATATGAGCCCCCAGAGATCAGCCTGG 985
QY      234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 252
Db      986 ACCGGTCACGGCCACCCACAGCCCGCAGTCAAAAGCTGCTCAACCATCTCTTCCACAGTG 1045
QY      253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
Db      1046 CCCAAGACTGAACACAGCGCTCTCATGATTAGATCTTATGATTCCTTGAGCCACACAGT 1105
QY      273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 292
Db      1106 AGCCGCTTGCAAAATCCAGGCACTGGCCAGATCCACTTGGCAGATTCCTCGAGACTC 1165
QY      293 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyIlePheLys 312
Db      1166 CTGTGCGACAGCTCCAACTCCAGCTCATCTCCAGTGGGAAGGCCAACCGGGAGTTCAAG 1225
QY      313 MetThrAspProAspGlnValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332
Db      1226 ATACGCGATCCGACAGAGTGGCCCGCGCTGGGGAGAGCGGAGAGCAACCAACACATG 1285
QY      333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352
Db      1286 AACTACAGATTAAGCTCAGCCGCGCGCTCCGTTACTACTATGACAAAGAACATCATGACCAAG 1345
QY      353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
Db      1346 GTCCATGGGAAAGCGCTACGCTTCAAGTTCTGACTCCAGGGGATCCCGACGGCCCTCCAG 1405
QY      373 ProHisProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 392
Db      1406 CCCCACCCCGGAGTCACTCTGTACAGATACCCCTTCAGACCTCCCGTACATGGGCTCC 1465
QY      393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412
Db      1466 TATCACGCCGCCACCAAGAAAGATGAACTTGTGGCCGCCACCCCTCCAGCCGCCGCG 1525
QY      413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIle 432
Db      1526 ACATCTTCAAGTGTTCCTGCTGCCCAAAACCCCTACTGGAATTCACCAACTGGGGGTATA 1585
QY      433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db      1586 TACCCCAACACTAGGCTCCCAACAGCCATATGCTTCTCATCTGGGCACTTACTAC 1642

RESULT 10
US-09-784-356-26
; Sequence 26, Application US/09784356
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis,

```

```

; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-784-356-26

Alignment Scores:
Pred. No.: 1,86e-160 Length: 3166
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 88.67% Mismatches: 24
Query Match: 89.28% Indels: 14
DB: 30 Gaps: 4

US-09-902-772-2 (1-451) x US-09-784-356-26 (1-3166)
QY 2 AlasethrlllelysgualaleuSerValSerclunspInserleupheglucys 21
DB 281 GCAGCTCATATCAAGAGAGCCCTTATCATGTTGATGAGAGACCAAGTCGTTGAGCTGT 340
QY 22 AlatyrglyserProhlsleualalystrhglumethrhalasethrSerSerglutyr 41
DB 341 GCTACGAGAGACGACACCTGCTAGACAGAGATGACCCGTCCTCTCCAGAGACTAT 400
QY 42 GlyglthrserlysmetSerProargValProglnglnasprtleusSerlnProPro 61
DB 401 GGACAGACTTCCAGAGAGACCCACGCGCTCCACAGAGAGATTGGCTGTCAACCCCA 460
QY 62 AlargValthrlllelysmetGlucysAsnProasnglnValasnglySerArgasner 81
DB 461 GCAGGCTCACATCAAAATGATGATGAACCTAGCCAGGTGATGCTCAAGAACTCT 520
QY 82 ProaspspcysSerValalalyglylysmetValserSerSerSeraspaasnValgly 101
DB 521 CCGATGAGATGAGTGTGGCCAAAGCGGGAGAAATGGTGGGAGCCAGACCGTTGGG 580
QY 102 MetasntylglyserlysmetglulysHisllleProProProasmetThrtrhAsn 121
DB 581 ATGACTACGCGCAGCTACATGAGAGAGAACACATGCCACCCCAACATGACACGAAAC 640
QY 122 GluArgArgValllleValProAlaAspProThrleuTrpSerThrAspHisValArgln 141
DB 641 GACCGCAGACTTCTGTCGACAGACAGATCCACCTATGAGTACAGACCATGTGCGGAG 700
QY 142 TrpleuglutrAlaVallysglutyrGlyleuProaspylaAspilleuLeuphegln 161
DB 701 TGGCTGAGAGTGGCGGTGAAGAATATGGCTTCCAGACGCTCAACATCTGTATTCCAG 760
QY 162 AsnllAspGlylysgluLeuGlysmetThrlyAspAspAspAspAspAspAspAsp 181
DB 761 AACATGATGAGAGAGAGAGTGTCAAGATGACCAAGAGACACTTCCAGAGAGCTACCC 820
QY 182 SerTyrAsnAlaAspilleuLeuSerHisllleuHisTyrleuArglnArglnAlaThr 201
DB 821 AGCTACAGCCGACATCTTCTCTCATCATCTCCACTCAGAGAGACTCCTCTTCCA 880
QY 202 PheillepheProasnthrSerValTyrProgluAlaThrGlnArgllethrThrArgPro 221
DB 881 CATTTG-----ACTTCA-----GATGATGTGATTAAGCCTTACAAACTCTCCA 925
QY 222 -----AspleuProTyrGluGlnAlaArgArgSerAlaTrp 233

```

```

DB 926 CGGTAAATGATCTTGAAACACAGATTTACATATGAGACCCCCAGAGATAGCCTGG 985
QY 234 ThrSerHisSerHisProthr-----GlnserlyAlaThrGlnProSerSerThrVal 252
DB 986 ACCGGTCAGCGGCACCCACAGCCAGGTGAGAAAGCTGCTCAACATCTCTTCCACAGTG 1045
QY 253 ProlysthrGlnAspGlnArgProGlnLeuAspProTyrGlnlleuLeuglyProthrSer 272
DB 1046 CCCAAACGTGAACACAGCGCTCTCAGTTAGATCTTATCAGATCTTGGACCAACAGT 1105
QY 273 SerArgLeuAlaAsnProGlyserGlyGlnlleGlnleuThrPbInpheuLeuGlnLeu 292
DB 1106 AGCCGCTTGCATATCCAGGACAGTGGCCAGATCCAGATCTTGGCATTTCTCTCGAGCTC 1165
QY 293 LeuSerAspSerSerAsnSerAsnCyslleThrTrpGluGlyThrAsnGlyGluPheLys 312
DB 1166 CTGTCGAGACAGCTCCAACTCCACTGCTGATCCTCGGGAAGGCACCAAGGGAGTTCAAG 1225
QY 313 MetThrAspProAspGluValAlaArgArgTrpGlyGluArglySerLysProasmet 332
DB 1226 ATGACGAGATCCGACAGAGGTGGCCGCGCTGGGAGAGCGGAAAGACCAACCAACATG 1285
QY 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnlleuMetThrLys 352
DB 1286 AACTACGATTAAGCTCAGCGCGCGCTCCCTACTACTATGACAAAGACATCATGACCAAG 1345
QY 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlylleAlaGlnAlaLeuGln 372
DB 1346 GTCCATGAGAGAGGCGTACGCGCTACAGTTGATTCACAGGATGCGCCAGCGCTCCAG 1405
QY 373 ProhlsProProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 392
DB 1406 CCCACCCCGGAGATCATCTGTCAAGATCCCTCAGACCTCCGATCATGAGGCTCC 1465
QY 393 TyrHisAlaHisProGlnLysMetAsnPhleValAlaProhlsProProAlaLeuProVal 412
DB 1466 TATCAGCCCGCACCAAGAGATGATGTTGTTGCGCCACCCACCTCAGCGCTCCAGCGTG 1525
QY 413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyle 432
DB 1526 ACATCTTCCAGTTTTTTTGTGTGCCCAAAACCATCTGGAATTCACCAACTGGGGGTATA 1585
QY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisleuGlyThrTyrTyr 451
DB 1586 TACCCCAACACTAGGCTCCCGCACACCATATGCTTCTCATCTGTGGCACTTACTAC 1642

RESULT 11
US-10-021-660-26
; Sequence 26, Application US/10021660
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-021-660-26

Alignment Scores:

```

Pred. No.: 1,86e-160 Length: 3166  
 Score: 2182.00 Matches: 407  
 Percent Similarity: 91.72% Conservative: 14  
 Best Local Similarity: 88.67% Mismatches: 24  
 Query Match: 89.28% Indels: 14  
 DB: 38 Gaps: 4

US-09-902-772-2 (1-451) x US-10-021-660-26 (1-3166)

QY 2 AlaSerThrIleIysGluAlaLeuSerValValSerGluAspGlnSerIleuPheGluCys 21  
 DB 281 GCACCTATATCAAGGAGGAGCCCTTATCACTGTGAGTGAGGACAGTCCTTTGAGAGT 340  
 QY 22 AlaTyrGlySerProHisIleuAlaIlyThrGluMetThrAlaSerSerSerGlnTyr 41  
 DB 341 GCCTACGAGACGCCACACCTGGCTTAAGACAGATGACGCCGCTCTCTCCAGGACAT 400  
 QY 42 GlyIleThrSerIysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 61  
 DB 401 GGACAGACTTCCAGATGAGCCACGCGTCCCTCAGAGATGGCTGTCTCAACCCCA 460  
 QY 62 AlaArgValThrIleIysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
 DB 461 GCCAGGGTCACCATCAAAATGGAATGTAACTTACCCAGGTGAATGGCTCAAGAGACTCT 520  
 QY 82 ProAspAspCysSerValAlaIlyGlyIlyMetValSerSerSerAspAsnValGly 101  
 DB 521 CCTATGATCATCTAGTGGCCAAAGCGGAGAGATGGTGGGACGCCAGACACCGTTGG 580  
 QY 102 MetAsnTyrGlySerTyrMetGluGluIlyHisIleProProProAsnMetThrAsn 121  
 DB 581 ATGAACTACAGCGACCTACATGAGAGAGACATGCGACCCCAACATGACCAACG 640  
 QY 122 GluArgThrValIleValProAlaAspProThrIleuTyrSerThrAspHisValArgGln 141  
 DB 641 GAGGCGAGATTATCGTCCAGAGATCTTACGCTATGAGATACAGACAGTGGCGAG 700  
 QY 142 TyrPleuGluTyrPalaValIlyGluTyrGlyLeuProAspValAspIleLeuPheGln 161  
 DB 701 TGGCTGAGTGGCGGCTGAAGATATGCTTCACAGCGTCACATCTTTATTTCCAG 760  
 QY 162 AsnIleAspGlyIysGluLeuCysIysMetThrIysAspAspPheGlnArgLeuThrPro 181  
 DB 761 AACATCGATGGAGAGAGACTGTGCAAGATGACCAAGAGCACTTCCAGAGCTCACCCC 820  
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluArgGlyAlaThr 201  
 DB 821 AGCTACACAGCGCGACATCTTCTCTCACATCTCCACTACAGAGACACTCTTCCA 880  
 QY 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgPro 221  
 DB 881 CATTTG-----ACTTCA-----GATGATGTGATTAAAGCCTTCAAAACTCTTCA 925  
 QY 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTyr 233  
 DB 926 CGGTAATGATGCTAGAAACACAGATTACCATATAGAGCCCAAGAGATGAGCCCTGG 985  
 QY 234 ThrSerHisSerHisProThr-----GlnSerIysAlaThrGlnProSerSerThrVal 252  
 DB 986 ACCGGTACAGGCGACCCACCCAGCGAGTCGAAAGCTGCTCAACCATCTCTTCCACAGT 1045  
 QY 253 ProIysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272  
 DB 1046 CCCAAACTGAAGACAGAGCTCTCTCAGTTAGATCTTATGATCTTGGGACCAACAGT 1105  
 QY 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuGluLeu 292  
 DB 1106 AGCGGCTTGCAGAAATCCAGAGAGTGGCCAGATCCAGTTTGGAGTTCTCTGGAGCTC 1165  
 QY 293 LeuSerAspSerSerAsnSerAsnGlyIleThrTyrGluGlyThrAsnGlyGluPheIys 312  
 DB 1166 CTGTGCGACACTCCAACTCAGCTGATCAGCTGAGGAGGAGGAGGAGGAGTTCAG 1225

QY 313 MetThrAspProAspGluValAlaArgArgTyrGlyIleArgIysSerIysProAsnMet 332  
 DB 1226 ATGAGGATGCCGAGAGAGTGGCGGCTGGGAGAGCGGAGAGCAAAACCAACATG 1285  
 QY 333 AsnTyrAspIysLeuSerArgAlaLeuArgTyrTyrTyrAspIysAsnIleMetThrIys 352  
 DB 1286 AACTACGATTAAGCTCAGCGCGGCGCTCGTACTACTATGATGACAGAAATCATGACCAAG 1345  
 QY 353 ValHisGlyIysArgTyrAlaTyrIlyIysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372  
 DB 1346 GTCCATGGGAGAGCGCTACGCTACAGATTCGACTTCCACGGGATGCCAGGCCCTCCAG 1405  
 QY 373 ProHisProProGluSerSerMetTyrIlyTyrProSerAspLeuProTyrMetSerSer 392  
 DB 1406 CCCAGCCCCCGGAGTATCTCTGTACAGATACCCCTCAGACCTCCGTATATGGGCTCC 1465  
 QY 393 TyrHisAlaHisProGlnIlyMetAsnPheValAlaProHisProProAlaLeuProVal 412  
 DB 1466 TATCAGCGCCCAACACAGAAATGAACTTTGTGGCGGCCCACTCCAGCCCTCCGCTG 1525  
 QY 413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGlyIle 432  
 DB 1526 ACATCTTCCAGTTTTTTTGTGTCGCCCAACCCATACCTGAATTCACCAACTGGGGATTA 1585  
 QY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 DB 1586 TACCCCAACACTAGGCTCCCAACAGCATATGCTTCTCATCTGGGACTTACTAC 1642

RESULT 12  
 US-10-205-823-98  
 ; Sequence 98, Application US/10205823  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schlegel, Robert  
 ; APPLICANT: Monahan, John E.  
 ; APPLICANT: Endege, Wilson O.  
 ; APPLICANT: Gannavarapu, Manjula  
 ; APPLICANT: Gorbacheva, Bella  
 ; APPLICANT: Hoersch, Sebastian  
 ; APPLICANT: Kamatkar, Shubhangi  
 ; APPLICANT: Womsey, Angela M.  
 ; APPLICANT: Glatt, Karen  
 ; APPLICANT: Zhao, Xumei  
 ; APPLICANT: Anderson, Dustin  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
 ; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; FILE REFERENCE: MRI-044  
 ; CURRENT APPLICATION NUMBER: US/10/205,823  
 ; CURRENT FILING DATE: 2002-07-25  
 ; PRIOR APPLICATION NUMBER: 60/307,982  
 ; PRIOR FILING DATE: 2001-07-25  
 ; PRIOR APPLICATION NUMBER: 60/314,356  
 ; PRIOR FILING DATE: 2001-08-22  
 ; PRIOR APPLICATION NUMBER: 60/325,020  
 ; PRIOR FILING DATE: 2001-09-25  
 ; PRIOR APPLICATION NUMBER: 60/341,746  
 ; PRIOR FILING DATE: 2001-12-12  
 ; PRIOR APPLICATION NUMBER: 60/362,158  
 ; PRIOR FILING DATE: 2002-03-05  
 ; NUMBER OF SEQ ID NOS: 455  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 98  
 ; LENGTH: 3166  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-205-823-98

Alignment Scores:  
 Pred. No.: 1,86e-160 Length: 3166  
 Score: 2182.00 Matches: 407  
 Percent Similarity: 91.72% Conservative: 14  
 Best Local Similarity: 88.67% Mismatches: 24  
 Query Match: 89.28% Indels: 14







QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140  
 Db 467 AATGAGGCGAGAGTGCCTCCCGAGATGCTCTGGGAGACAGACAGCATGCTCGA 526  
 QY 141 GluTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPhe 160  
 Db 527 CAGGGCTGGGTGGGGGTGAAGAAATATGGCTCTCGATGGGACGCTTACTATT 586  
 QY 161 GluAsnIleAspGlyLysGluLeuValLysMetThrLysAspAspPheGluArgLeuThr 180  
 Db 587 CAGATATCGATGGGAGAGCTGTGCAGATGCAAGAGTGAAGTGAAGTCTTCAGCGGCTCAG 646  
 QY 181 ProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisThrLeuArgGluArg----- 198  
 Db 647 CCGAGCTACATGCCGACATTTCTCTCATATCCACATCCGACAGAGACACCCCTT 706  
 QY 198 ----- 198  
 Db 707 CCACATCTGACTTCGATGACGTTGATTAAGCTTTACAAAACCTCCACGGTTAATGCA 766  
 QY 199 ----- 213  
 Db 767 GCCAGAAACACAGGGGGTGCAGCTTTATTTCCCAATCTTCAGTATATCCGGAAGCT 826  
 QY 214 ThrGluArgGluThrThrArgProAspLeuProTyrGluGluAlaArgSerAlaTrp 233  
 Db 827 ACGCAAGAAATTAACACTAG----- 847  
 QY 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 253  
 Db 847 ----- 847  
 QY 254 LysThrGluAspGluArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273  
 Db 847 ----- 847  
 QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeuLeu 293  
 Db 848 ----- 895  
 QY 294 SerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLysPheLysMet 313  
 Db 896 TCAGACAGCTCCAACTCAATGCTCATCTGCGGAGAGCAACAAAGGGAGTTCAAGATG 955  
 QY 314 ThrAspProAspGluValAlaArgArgTrpGlyLysArgLysSerLysProAsnMetAsn 333  
 Db 956 ACAGACCCGAGACGAGTGGCTGGGGGAGAGAGCAAGCAAGCCCAACATGAAAC 1015  
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353  
 Db 1016 TATGACAAAGCTCAGCCGCCCTCGCTACTACTACGACAAAACATCATGACCAAGTG 1075  
 QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373  
 Db 1076 CACGGGAGGCTACAGCTCAAGTTGACTTCCACGGAGATGGCCAGGCCCTGCACAGCC 1135  
 QY 374 HisProProGlnUserSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393  
 Db 1136 CACCTCTGATGCTGCTGCTGATACAGTACCCCTCCAGCTGCGCATCATGAGGCTCTAT 1195  
 QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413  
 Db 1196 CAGCCCAACCCGAGAGATGAATTTGTGTCCGCCACCTCCCGCTGCCAGTGCACA 1255  
 QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyr 433  
 Db 1256 TCTTCCAGTTCTTCTTCCCGGACCAACCATACGGAATTCACCGAGATGGGGGAGATCTAC 1315  
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 Db 1316 CCGAGACACTAGGCTCCAGCCAGCCATATGCTCTCAGCTGGGACACTTACTAC 1369

US-09-442-384A-475  
 : Sequence 475, Application US/09442384A  
 : GENERAL INFORMATION:  
 : APPLICANT: Chenchik, Alex  
 : APPLICANT: Lukashiev, Malvey  
 : TITLE OF INVENTION: Hematology/Immunology Array  
 : FILE REFERENCE: CLON-006CIP15  
 : CURRENT APPLICATION NUMBER: US/09/442,384A  
 : PRIORITY FILING DATE: 1999-11-17  
 : PRIORITY FILING DATE: 1998-03-31  
 : NUMBER OF SEQ ID NOS: 830  
 : SOFTWARE: FASTSEQ for Windows Version 4.0  
 : SEQ ID NO 475  
 : LENGTH: 3126  
 : TYPE: DNA  
 : ORGANISM: homo sapiens  
 : US-09-442-384A-475  
 Alignment Scores:  
 Pred. No.: 2,72e-131 Length: 3126  
 Score: 1810.00 Matches: 336  
 Percent Similarity: 90.34% Conservative: 10  
 Best Local Similarity: 87.73% Mismatches: 23  
 Query Match: 74.06% Indels: 14  
 DB: 18 Gaps: 4  
 US-09-902-772-2 (1-451) x US-09-442-384A-475 (1-3126)  
 QY 78 SerArgAsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSer 97  
 Db 150 TCAAGGAATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 209  
 QY 98 AspAsnValGlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsn 117  
 Db 210 GACACCGTTGGGATGAATCAAGCAGCTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 269  
 QY 118 MetThrAsnGlnValArgValIleValProAlaAspProThrLeuTrpSerThrAsp 137  
 Db 270 ATGACCCGAAACGAGCCGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 329  
 QY 138 HisValArgGlnTrpLeuGluTrpAlaValLysGluLysGluLeuProAspValAspIle 157  
 Db 330 CATGTGGGAGAGTGGCTGGAGTGGGGGTGAAGAAATATGAGCTTCCAGAGCTCAACATC 389  
 QY 158 LeuLeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGln 177  
 Db 390 TTGTTATTCAGAAATCGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 449  
 QY 178 ArgLeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisThrLeuArgGlu 197  
 Db 450 AGGCTCACCCCGAGCTCAAGAGCCGAGATCTTCTCTCATATCCAGTCTCAGAGAG 509  
 QY 198 ArgGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIle 217  
 Db 510 ACTGCTCTCCACATTTG-----ACTTCA-----GATGATGTGATTAAGCCTTA 554  
 QY 218 ThrThrArgPro-----AspLeuProTyrGluGlnAlaArg 229  
 Db 555 CAAAACTCTCCAGAGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614  
 QY 230 ArgSerAlaTrpThrSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSer 248  
 Db 615 AGATCAGCTGGAGACGGGTTCACGGCCACCCACGCCAGTGCAGAAAGTGTCTCAACATCT 674  
 QY 249 SerSerThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeu 268  
 Db 675 CTTCCAGACTGCCCAAAATCGAAGACCAAGCGTCTCAGTTAATGATGATGATGATGATGATGAT 734  
 QY 269 GlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPhe 288  
 Db 735 GGACCAACAGTAGCGGCTTGGCAAAATCCAGGAGTGGCCAGATCCAGCTTGGGAGTTTC 794



Wed Jul 30 13:29:52 2003

us-09-902-772-2.rnp

Page 15

Search completed: July 28, 2003, 09:24:09  
Job time : 3002.3 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:30:39 ; Search time 1030.65 Seconds  
(without alignments)  
3233.216 Million cell updates/sec

Title: US-09-902-772-2  
Perfect score: 2444  
Sequence: 1 MASTIKALSVESDLSLE.....IYPNTRLPAAHPSHLCTTY 451

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 8407786 segs, 3694357880 residues  
Total number of hits satisfying chosen parameters: 16815572

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framet\_p2n.model -DEV=x1h  
-O=/cgn2\_1/USPTO.spool/US09902772/runat\_23072003\_093658\_14943/app\_query.fasta.1.1294  
-DB=Pending\_Patents\_NA\_New -QFMT=fastcap -SUFFIX=trpna -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cd1  
-LIST=45 -DOCALLIG=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US09902772.cgn1.1.325 @runat\_23072003\_093658\_14943 -NCPU=6 -ICPU=3  
-NO\_MMAP -IAREQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending\_Patents\_NA\_New:\*  
2: /cgn2\_6/ptodata/2/pna/PCN\_NEW\_COMB.seq4:\*  
3: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq4:\*  
4: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq4:\*  
5: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq4:\*  
6: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq4:\*  
7: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq4:\*  
8: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq4:\*  
9: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq4:\*  
10: /cgn2\_6/ptodata/2/pna/US12\_NEW\_COMB.seq4:\*  
11: /cgn2\_6/ptodata/2/pna/US13\_NEW\_COMB.seq4:\*  
12: /cgn2\_6/ptodata/2/pna/US14\_NEW\_COMB.seq4:\*  
13: /cgn2\_6/ptodata/2/pna/US15\_NEW\_COMB.seq4:\*  
14: /cgn2\_6/ptodata/2/pna/US16\_NEW\_COMB.seq4:\*  
15: /cgn2\_6/ptodata/2/pna/US17\_NEW\_COMB.seq4:\*  
16: /cgn2\_6/ptodata/2/pna/US18\_NEW\_COMB.seq4:\*  
17: /cgn2\_6/ptodata/2/pna/US19\_NEW\_COMB.seq4:\*  
18: /cgn2\_6/ptodata/2/pna/US20\_NEW\_COMB.seq4:\*  
19: /cgn2\_6/ptodata/2/pna/US21\_NEW\_COMB.seq4:\*  
20: /cgn2\_6/ptodata/2/pna/US22\_NEW\_COMB.seq4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result

Query

No.	Score	Match	Length	DB	ID	Description
1	2298	94.0	1930	15	US-10-170-235-28132	Sequence 28132, A
2	2298	94.0	1930	15	US-60-452-680-11396	Sequence 11396, A
3	2197	89.9	3294	15	US-10-170-235-28519	Sequence 28519, A
4	2197	89.9	3294	18	US-60-452-680-11397	Sequence 11397, A
5	2182	89.3	1389	2	PCT-US02-04915-41	Sequence 41, Appl
6	2182	89.3	1389	14	US-10-211-662-41	Sequence 41, Appl
7	2182	89.3	1389	15	US-10-087-192-1127	Sequence 1127, Ap
8	2182	89.3	1389	15	US-10-126-052A-330	Sequence 330, App
9	2057	84.2	1297	15	US-10-087-192-1124	Sequence 1124, App
10	1991	81.5	2172	14	US-10-144-771-355	Sequence 355, App
11	1624.5	66.5	1359	2	PCT-US02-41414-1210	Sequence 1210, App
12	1624.5	66.5	2957	2	PCT-US02-41414-1209	Sequence 1209, App
13	1624.5	66.5	2957	14	US-10-342-887-710	Sequence 710, App
14	1613.5	66.0	1359	2	PCT-US02-41414-1207	Sequence 1207, App
15	1613.5	66.0	1729	14	US-10-144-771-10845	Sequence 10845, A
16	1613.5	66.0	1729	14	US-10-144-771-10845	Sequence 34391, A
17	1610.5	65.9	2769	15	US-10-170-235-34391	Sequence 4688, Ap
18	1607.5	65.8	2172	11	US-09-949-016-4688	Sequence 1126, Ap
19	956.5	39.1	212231	15	US-10-087-192-1126	Sequence 1126, Ap
20	956.5	39.1	8059021	9	US-09-947-914-53	Sequence 53, Appl
21	949	38.8	69090	15	US-10-087-192-1123	Sequence 1123, Ap
22	948	38.8	567	14	US-10-203-138A-6819	Sequence 6819, Ap
23	863	35.3	473	14	US-10-203-138A-8888	Sequence 8888, Ap
24	833.5	34.1	57726	11	US-09-949-016-16430	Sequence 16430, A
25	833.5	34.1	73725	2	PCT-US02-41414-1208	Sequence 1208, Ap
26	813.5	33.3	72732	2	PCT-US02-41414-1205	Sequence 1205, Ap
27	654	26.8	533	10	US-09-513-999C-1340	Sequence 1340, Ap
28	654	26.8	533	12	US-09-513-999C-1340	Sequence 1340, Ap
29	650	26.6	420	14	US-10-203-138A-3751	Sequence 3751, Ap
30	650	26.6	454	14	US-10-203-138A-1690	Sequence 1690, Ap
31	511	20.9	1752	15	US-10-219-051B-13871	Sequence 13871, A
32	509.5	20.8	1876	15	US-10-170-235-24888	Sequence 24888, A
33	509.5	20.8	1876	18	US-60-452-680-295	Sequence 295, App
34	509.5	20.8	1901	15	US-10-219-051B-13873	Sequence 13873, A
35	499	20.4	1165	15	US-10-170-235-24404	Sequence 24404, A
36	498	20.4	318	18	US-60-452-680-296	Sequence 296, App
37	498	20.4	318	10	US-09-513-999C-25950	Sequence 25950, A
38	498	20.4	318	12	US-10-513-999C-25950	Sequence 25950, A
39	482	19.7	1581	15	US-10-170-235-34170	Sequence 34170, A
40	482	19.7	1581	19	US-60-455-444-673	Sequence 673, App
41	482	19.7	1581	19	US-60-455-444-673	Sequence 673, App
42	465	19.0	443	14	US-10-144-771-21735	Sequence 21735, A
43	460.5	18.8	2422	15	US-10-170-235-34449	Sequence 34449, A
44	460.5	18.8	2422	18	US-60-452-680-11304	Sequence 11304, A
45	460	18.8	1965	14	US-10-144-771-20068	Sequence 20068, A

## ALIGNMENTS

RESULT 1  
US-10-170-235-28132  
Sequence 28132, Application US/10170235  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig  
TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU  
TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF  
FILE REFERENCE: C1001380  
CURRENT APPLICATION NUMBER: US/10/170,235  
NUMBER OF SEQ ID NOS: 42514  
SEQ ID NO 28132  
LENGTH: 1930  
TYPE: DNA  
ORGANISM: HUMAN  
US-10-170-235-28132

Alignment Scores:  
Pred. No.: 2,47e-256  
Score: 2298.00  
Percent Similarity: 91.21%  
Best Local Similarity: 89.12%  
Length: 1930  
Matches: 426  
Conservative: 10  
Mismatch: 14

Query Match: 94.03% Indels: 28  
 DB: 15 Gaps: 2  
 US-09-902-772-2 (1-451) x US-10-170-235-28132 (1-1930)

QY 2 AlaserthrilleysguAlaLeuSerValValSerGluspgInserLeupheglucys 21  
 DB 25 GCACCTCATATCAAGGAGGAGCCTTATCAGTTGTGAGTACAGGACCAGTGTGTTGAGTGT 84  
 QY 22 AlatyrglySerProHIsleuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41  
 DB 85 GCCTACGAGAACGCCACACCTGGCTTAAGACAGATGACGGCTCTCTCTCCAGCGACTAT 144  
 QY 42 GlyInThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61  
 DB 145 GGACAGACTTCCAAAGATGAGCCCGCGGTCCCTCAGCAGATTTGGCTGTCTCAACCCCA 204  
 QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
 DB 205 GCCAGGGTCACCATCAAAATGGAATGTAAACCTTACAGTGAATGGCTCAAGGACTCT 264  
 QY 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnValGly 101  
 DB 265 CCTGATGAATGCACTGTGGCCAAAGCGGGAAGATGGGGGACACGACACCGCTGGG 324  
 QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121  
 DB 325 ATGAACACAGCGCAGCTCATGTGAGAGAGACATGCCACCCCAACATACACACGAC 384  
 QY 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141  
 DB 385 GAGCGCAGAGTTATCGCGCCAGCAGATCTACGCTATGAGTACAGACATGTGGCGAG 444  
 QY 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161  
 DB 445 TGGCTGAGTGGGGGGGGAAGATATGGCTTCCAGACGTCACATCTGTTATTCAG 504  
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181  
 DB 505 AACATCATGTGGAAGGAGTGTGCATGATGACCAAGACGACTCCAGAGGCTCACCC 564  
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluArg----- 198  
 DB 565 AGCTACACAGCGCAGATCTTCTCTCATCTCCATCCAGACAGAGACTCTCTTCCA 624  
 QY 198 ----- 198  
 DB 625 CATTTGACTTCAGATGATGTGATAAAGCCTTACAAACTCTCCAGCGTTAATGCATGCT 684  
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 214  
 DB 685 AGAAACACAGGGGTGACGCTTTTATTTCCCAATACCTTCAGATATCCGAGGACTAG 744  
 QY 215 GluArgIleThrThrArgProAspLeuProTyrGlnIleAlaArgSerAlaTrpThr 234  
 DB 745 CAAGAATTTCAACTAGGCGCAGATTTACATATAGCCCCCGCAGAGATCAGCTTGAGC 804  
 QY 235 SerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThrValPro 253  
 DB 805 GGTACAGGCCACCCACGCCCCAGTCCAAAGCTCTTACACATCTCTCTCCACAGTGGCC 864  
 QY 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273  
 DB 865 AAAAAGTGAAGACCGCTCTCAGTTAGATCCATTATCAGATTTTGAGACCAACAGTAGC 924  
 QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeu 293  
 DB 925 CGCCTTCAATCCAGGCAAGTGGCCAGATCCAGTTTGGAGATTCCTCCGAGACTCTCG 984  
 QY 294 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLupheLysMet 313  
 DB 985 TCGGACAGCTCCAACTCAGCTGATCAGCTGGGAGAGCAACAGGGGAGTTCAAGATG 1044

QY 314 ThrAspProAspGluValAlaArgArgTrpGlyLupheLysSerLysProAsnMetAsn 333  
 DB 1045 ACGGATCCCGCAGAGGAGGCGCGCTGGGGAGAGCGGAAGACCAACCATGATGAC 1104  
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353  
 DB 1105 TACGATAGCTCACCGCGCGCTCGTTACATATGACAAGAACATCATGACCAAGGTC 1164  
 QY 354 HisLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373  
 DB 1165 CATGGGAAGCGCTACGGCTTACAACTTGCACTTCCACGGAGATGCCCGCTCCAGCC 1224  
 QY 374 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393  
 DB 1225 CACCCCGGAGTATCTGTACAGAACCCCTCAGACCTCCGCTATGATGGCTCTAT 1284  
 QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413  
 DB 1285 CACGCCACCCACAGAGATGAACTTGTGGCGCCACCCCTCCAGCCCTCCCGTGACA 1344  
 QY 414 SerSerSerPhePheAlaProAsnProTyrTrpAsnSerProThrGlyLysTyr 433  
 DB 1345 TCTTCCAGTTTTTTTGTCTGCCCAACCCATCTGGAATTCACCACTGGGGGTATATAC 1404  
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 DB 1405 CCCAACACTAGGCGTCCCGCACAGCATATGCTTCTCATGTGGCCACTTACTAC 1458

RESULT 2  
 US-60-452-680-11396  
 : Sequence 11396, Application US/60452680  
 : GENERAL INFORMATION:  
 : APPLICANT: CARGILL, Michele  
 : TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 : FILE REFERENCE: CLO01450  
 : CURRENT APPLICATION NUMBER: US/60/452,680  
 : NUMBER OF SEQ ID NOS: 116213  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 11396  
 : LENGTH: 1930  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-60-452-680-11396

Alignment Scores:  
 Pred. No.: 2,47e-256 Length: 1930  
 Score: 2298.00 Matches: 426  
 Percent Similarity: 91.21% Conservative: 10  
 Best Local Similarity: 89.12% Mismatches: 14  
 Query Match: 94.03% Indels: 28  
 DB: 18 Gaps: 2

US-09-902-772-2 (1-451) x US-60-452-680-11396 (1-1930)

QY 2 AlaserthrilleysguAlaLeuSerValValSerGluspgInserLeupheglucys 21  
 DB 25 GCACCTCATATCAAGGAGGAGCCTTATCAGTTGTGAGTACAGGACCAGTGTGTTGAGTGT 84  
 QY 22 AlatyrglySerProHIsleuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41  
 DB 85 GCCTACGAGAACGCCACACCTGGCTTAAGACAGATGACGGCTCTCTCTCCAGCGACTAT 144  
 QY 42 GlyInThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61  
 DB 145 GGACAGACTTCCAAAGATGAGCCCGCGGTCCCTCAGCAGATTTGGCTGTCTCAACCCCA 204  
 QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
 DB 205 GCCAGGGTCACCATCAAAATGGAATGTAAACCTTACAGTGAATGGCTCAAGGACTCT 264

QY 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101  
 Db 265 CCGATGATGATGAGTGGCCAAAGGGGGAAGATGGTGGGACAGCCAGACACCGTTGGG 324  
 QY 102 MetAsnTyrGlySerTyrMetGluGluHisIleProProProAsnMetThrAsn 121  
 Db 325 ATGAACTACGCGACGCTCATGGAGAGACACATGCGACCCCAACATGACACAGCAGAC 384  
 QY 122 GluArgArgValIleValProAlaAspProThrIleuThrSerThrAspHisValArgGln 141  
 Db 385 GAGCGGAGAGTATCGCCAGCAGATCTACGCTATGAGATGACAGCCATGTGGCGAG 444  
 QY 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161  
 Db 445 TGGCTGAGTGGGGGGTGAAGAATATGGCTTCCAGATGCAACATCTGTATTTCAG 504  
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181  
 Db 505 AACATCGATGGAGGAGAACTGTCAGATGACCAAGAGACACATCCAGAGGCTCACCCCC 564  
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrIleuArgGluArg 198  
 Db 565 AGCTACACACCCGACATCTTCTCTCATCTCCACTACCTCAGAGAGACTCCCTTCCA 624  
 QY 198 ----- 198  
 Db 625 CATTTGACTTCAGATGATGTGATTAAGCCTTACAAAACCTCCACGAGTTAATCATGCT 684  
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 214  
 Db 685 AGAAGACAGAGGGGTGAGCTTTATTTATTTCCAAATCTTCAGATATATCTGAAAGCTACG 744  
 QY 215 GluArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrThr 234  
 Db 745 CAAGAGATTTACACTAGGCCAGATTTACATATGAGCCCGCCAGAGATACGCTTGACCC 804  
 QY 235 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 253  
 Db 805 GGTACAGCGCCAGCCAGCCAGCCAGTGCAGAAAGCTGCTCAACATCTCTTCCACAGTCC 864  
 QY 254 LysThrGluAspGlnArgProGlnIleuAspProTyrGlnIleuGlyProThrSerSer 273  
 Db 865 AAAACGTAAAGACAGCTCTCTCAGTTAGATCTTATGATTCCTTGAGCACCAACAAGTACG 924  
 QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheIleuLeuGluLeuLeu 293  
 Db 925 CCGCTTGCAATTCAGGACGATGGCCAGATCCAGCTGTGGCATGTTCTCTGGAGCTCTCG 984  
 QY 294 SerAspSerSerAsnSerAsnCysIleThrTrpGluGluLysThrAsnGlyLysPheLysMet 313  
 Db 985 TCGGACAGCTCCAACTCCAGCTGATCAGCTGGGAAAGGCACCAAGGGAGTTCAAGATG 1044  
 QY 314 ThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysProAsnMetAsn 333  
 Db 1045 ACGGATCCACAGAGGTGGCCGCGCTGGGAGAGGAGGAAAGCAAAACCAACATCAAC 1104  
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353  
 Db 1105 TAGGATAGCTCACCGCCGCCCTCCGTTACTACTATGACAAAGACATCAAGACCAAGTGC 1164  
 QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373  
 Db 1165 CAGGGAGAGCGGTACGCTCAAGTTGCTCAGGAGGATCGCCAGGCGCTCCAGCC 1224  
 QY 374 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393  
 Db 1225 CACCCCGGAGTATCTCTGTACAGTACCTCCAGACCTCCGTTCAAGGCGCTCTAT 1284  
 QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413  
 Db 1285 CAGGCCACCAAGAGATGAACTTGTGGCGCCACCCCTCCAGGCGCTCCCGGTACA 1344  
 QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 433

Db 1345 TCTTCCAGTATTTTGTGCCCCCAACCCACTAGTGAATTCACCACTGGGGATATATAC 1404  
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisIleuGlyThrTyrTyr 451  
 Db 1405 CCCAACACTAGGCTCCACACAGCATATGCTTTCATCTGCGGCACTTACTAC 1458  
 RESULT 3  
 US-10-170-235-28519  
 ; Sequence 28519, Application US/10170235  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig  
 ; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU  
 ; FILE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF  
 ; FILE REFERENCE: C1001380  
 ; CURRENT APPLICATION NUMBER: US/10/170, 235  
 ; NUMBER OF SEQ ID NOS: 42514  
 ; SEQ ID NO 28519  
 ; LENGTH: 3294  
 ; TYPE: DNA  
 ; ORGANISM: HUMAN  
 US-10-170-235-28519  
 Alignment Scores:  
 Pred. No.: 3,92e-244 Length: 3294  
 Score: 2197.00 Matches: 410  
 Percent Similarity: 91.96% Conservative: 13  
 Best Local Similarity: 89.13% Mismatches: 23  
 Query Match: 89.89% Indels: 14  
 DB: Gaps: 4  
 US-09-902-772-2 (1-451) x US-10-170-235-28519 (1-3294)  
 QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGln 20  
 Db 94 ATGGCCAGCACTATTAAAGAAAGCCTTATCAGTTGTGAGTAGAGACCAAGTGTGTTCAG 153  
 QY 21 CysAlaTyrGlySerProHisIleuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 Db 154 TGTGCTTACGGAAGCCACACTGGCTTAAGACAGAGTGAACGGGTCTCTCCAGCGAC 213  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 Db 214 TATGAGACAGACTTCCAAAGATGAGGCCAGCGGTCCCTCAGCAGATTTGGTCTCAACCC 273  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAlaAsnGlySerArgAsn 80  
 Db 274 CCAAGCCAGGCTCACATCAAAATGGAATGTAAACCTTACCCAGGTGAATGCTCAAGAGAC 333  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 Db 334 TCTCTGATTAATGCACTGTGGCCAAAGGGGGAAGTGGTGGGACACCCAGACACCGTT 393  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluHisIleProProProAsnMetThrThr 120  
 Db 394 GGGATGAACTACGCGACACTCATGAGAGAGAAGACATGCGACCCCAACATGACAGCAG 453  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuThrSerThrAspHisValArg 140  
 Db 454 AACGAGCGCAGAGTATCGTCCAGCAGATCTTACGCTATGAGTACAGACCAATGTGCGG 513  
 QY 141 GluTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160  
 Db 514 CAGTGGCTGAGTGGGGGTGAAGAATATGCGCTTCCAGACGTCACATCTGTATTC 573  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 574 CAGAACCTCATGGGAAGAGAACTGTGCAATATGACCAAGACACATCTCCAGAGGCTCAC 633  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluArgGlyAla 200  
 Db 634 CCAAGTACAAACCGCCAGATCTTCTCTCAGATCTCCACATCTCCATACCTCAGAGAGACTCTT 693

```

QY 201 ThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArg 220
Db 694 CCACATTTG-----ACTTCA-----GATGATGTGTTAAAGCCTTCAAAACTCT 738
QY 221 Pro-----AspleuProTyrGlnGlnAlaArgSerAla 232
Db 739 CCACGGTTAATGATGCTAGAAACACAGATTATACCATATGAGCCGCCAGAGATACAGCC 798
QY 233 TrpThrSerHisSerHisProThr---GlnSerLysAlaThrGlnInProSerSerSerThr 251
Db 799 TGGACCGGTCCAGGCCACCCACGCCGCCAGTGGAAAGTGGCTCAACCATCTCTTCCACA 858
QY 252 ValProLysThrGlnAspGlnArgProGlnInleuAspProTyrGlnIleleuGlyProThr 271
Db 859 GTGCCCAAAACTGAAAGACACAGCTCCAGTTAGATCCTTATCAGATTCTTGAGCAACA 918
QY 272 SerSerArgLeuAlaAsnProGlySerGlyInIleGlnLeuTrpGlnPheLeuLeuGlu 291
Db 919 AGTAGCCGCTTGGCAAAATCCAGGCAAGTGGCCAGATCCAGCTTGGCAGTTCTCTCGAG 978
QY 292 LeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGlnGluIleThrAsnGlyGluPhe 311
Db 979 CTCTGTGCGACACTCCACTCCAGCTGCTACCTGGAGGACACCCAGGAGGAGTTCC 1038
QY 312 LysMetThrAspProAspGluValAlaArgArgTrpGlyLysSerLysProAsn 331
Db 1039 AAGATGACGAGTCCCGACGAGAGTGGCCGCGCTGGGAGAGCGGAAGACAAACCCAC 1098
QY 332 MetAsnLysAspLysLeuSerArgAlaLeuArgTyrTyrTrpAspLysAsnIleMetThr 351
Db 1099 ATGAGACTACGATTAAGCTCAGCCGCGCTCGGTACTACATGACAAAGAAATCATGAGCC 1158
QY 352 LysValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeu 371
Db 1159 AAGTTCATGGGAGCGCTACAGCTGCTACAGTTGACATTCACGGGAGATCGCCAGCCCTC 1218
QY 372 GlnProHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSer 391
Db 1219 CAGCCCAACCCCGCGAGTATCTCTACAAAGTACCCTCAGACCTCCCTCATCATGGGC 1278
QY 392 SerTyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuPro 411
Db 1279 TCTCATACGCGCCACCCACAGAAAGATGACTTGTGGGCCCCACCTCCAGCCCTCCCC 1338
QY 412 ValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGly 431
Db 1339 GTGACATCTCCCAATTTTTCGTGCGCCCAAAACCCATACAGATTCACCAACTGGGGGT 1398
QY 432 IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1399 ATATACCCCAACACTAGGCTCCACACAGCATATGCTTCTCATCTGGGCACTTACTAC 1458

RESULT 4
US-60-452-680-11397
: Sequence 11397, Application US/60452680
: GENERAL INFORMATION:
: APPLICANT: CARBIL, Michele
: APPLICANT: GRUPE, Andrew
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001450
: CURRENT APPLICATION NUMBER: US/60/452,680
: NUMBER OF SEQ ID NOS: 116213
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11397
: LENGTH: 3294
: TYPE: DNA
: ORGANISM: Homo sapiens
US-60-452-680-11397
Alignment Scores:

```

```

Pred. No.: 3,92e-244 Length: 3294
Score: 2197.00 Matches: 410
Percent Similarity: 91.96% Conservative: 13
Best Local Similarity: 89.13% Mismatches: 23
Query Match: 89.89% Indels: 14
DB: 18 Gaps: 4
US-09-902-772-2 (1-451) x US-60-452-680-11397 (1-3294)

QY 1 MetLaserThrIleLysGlnAlaLeuSerValTyrSerGluAspGlnSerLeuPheGlu 20
Db 94 ATGGCCAGCACTATTAAAGAAAGCTTATCAGTTGTGATGAGACCAAGTGTGTTGAG 153
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerSerGlu 40
Db 154 TGTGCTACGGAAGACGCCACACCTGCTTAAGACAGATGACCGGCTCTCTCCAGGAC 213
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
Db 214 TATGAGACAGACTTCCAAAGATGAGCCACAGCGCTCCCTCAGCGAGTGGCTGTCAACC 273
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db 274 CCAGCCAGGGTCACATCAATAATGAAATGTAACCTGACCAAGTGAAATGGCTCAAGAAC 333
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db 334 TCTCTGATGAATGACATGTGGCCAAAGCGCGGAAGATGGTGGGACGCCAGACACCGTT 393
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
Db 394 GGGATGAACATACGGCACCTCATGTGAGAGAGAGATGCGCACCCCAACATGACACAG 453
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
Db 454 AAGCAGCGCAGAGTATTCGTGCCAGAGATCTTACGCTATGAGATACAGACATGTCCG 513
QY 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
Db 514 CAGTGGCTGGAGTGGCGGGTGAAGAAATATGAGCTTCCAAAGTCAACATCTGTATTTC 573
QY 161 GlnAsnIleAspGlyLysGluLeuLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 574 CAGAACTCATATGGGAAGAACTGTGCAAGATGCAAGAGCGCTTCCAGAGGCTCAC 633
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgLysGlyAla 200
Db 634 CCCAGCTACAAAGCGCCGACATCTCTCTCATATCTCCATACCTCAAGAGACATCTCT 693
QY 201 ThrPheIlePheProAsnThrSerValTyrTrpProGluAlaThrGlnArgIleThrThrArg 220
Db 694 CCACATTTG-----ACTTCA-----GATGATGTGTTAAAGCCTTCAAAACTCT 738
QY 221 Pro-----AspleuProTyrGlnGlnAlaArgSerAla 232
Db 739 CCACGGTTAATGATGCTAGAAACACAGATTATACCATATGAGCCGCCAGAGATACAGCC 798
QY 233 TrpThrSerHisSerHisProThr---GlnSerLysAlaThrGlnInProSerSerSerThr 251
Db 799 TGGACCGGTCCAGGCCACCCACGCCGCCAGTGGAAAGTGGCTCAACCATCTCTTCCACA 858
QY 252 ValProLysThrGlnAspGlnArgProGlnInleuAspProTyrGlnIleleuGlyProThr 271
Db 859 GTGCCCAAAACTGAAAGACACAGCTCCAGTTAGATCCTTATCAGATTCTTGAGCAACA 918
QY 272 SerSerArgLeuAlaAsnProGlySerGlyInIleGlnLeuTrpGlnPheLeuLeuGlu 291
Db 919 AGTAGCCGCTTGGCAAAATCCAGGCAAGTGGCCAGATCCAGCTTGGCAGTTCTCTCGAG 978
QY 292 LeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGlnGluIleThrAsnGlyGluPhe 311
Db 979 CTCTGTGCGACACTCCACTCCAGCTGCTACCTGGAGGAGACCAACAGGAGGAGTTCC 1038

```



QY 312 LysMetThrAspProaspGluValAlaIaArgArgTrpGluArgLysSerLysProasn 331  
 |||  
 Db 1039 AAGATGACGAGATCCCGAGAGAGTGGCCGGCGCTGGGAGAGCGGAAAGCAACCCAC 1098  
 QY 332 MetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThr 351  
 |||  
 Db 1099 ATGACCTACGATTAACCTCAGCGCGCCCTCCGTTACTACTATGACAAACATCATGTGACC 1158  
 QY 352 LysValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeu 371  
 |||  
 Db 1159 AAGGTCATGGGAACCGCTAGCGCCCTCAACAAGTTCCGACTTCCAGGGATGCCAGGCCCTC 1218  
 QY 372 GlnProHisProProGluSerSerMetLysTyrTyrProSerAspLeuProTyrMetSer 391  
 |||  
 Db 1219 CAGCCCAACCCCGGAGTCACTCTGTACAAGTACCCCTCAGACCTCCCGTACATGGGC 1278  
 QY 392 SerTyrHisAlaHisProGlnLysMetAsnPheValAlaIaProHisProProAlaLeuPro 411  
 |||  
 Db 1279 TCCTATACAGCCGCCACACAGAAAGTGAACCTTGTGGCGCCGCCACCTCCAGCCCTCCCC 1338  
 QY 412 ValThrSerSerSerPhePheAlaIaIaProAsnProTyrTrpAsnSerProThrGlyGly 431  
 |||  
 Db 1339 GTGACATCTTCAGTGTCTTTTGTCTGCCCCAACCATCTGGAATTCACCAACTGGGGGT 1398  
 QY 432 IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 |||  
 Db 1399 AATATACCCCAACACTAGGCTCCCAACGACATATGCTTCTCATCTGGGACCTTACTAC 1458

RESULT 5  
 PCT-US02-04915-41  
 ; Sequence 41, Application PC/TUS0204915  
 ; GENERAL INFORMATION:

; APPLICANT: Murray, Richard  
 ; APPLICANT: Glynn, Susan R.  
 ; APPLICANT: Watson, Susan R.  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Eos Biotechnology, Inc.  
 ; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and  
 ; FILE REFERENCE: 018501-006200PC  
 ; CURRENT APPLICATION NUMBER: PCT/US02/04915  
 ; CURRENT FILING DATE: 2002-02-14  
 ; PRIOR APPLICATION NUMBER: US 09/784,356  
 ; PRIOR FILING DATE: 2001-02-14  
 ; PRIOR APPLICATION NUMBER: US 09/791,390  
 ; PRIOR FILING DATE: 2001-02-22  
 ; PRIOR APPLICATION NUMBER: US 60/285,475  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: US 60/310,025  
 ; PRIOR FILING DATE: 2001-08-03  
 ; PRIOR APPLICATION NUMBER: US 60/350,666  
 ; PRIOR FILING DATE: 2001-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/334,244  
 ; PRIOR FILING DATE: 2001-11-29  
 ; NUMBER OF SEQ ID NOS: 230  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 41  
 ; LENGTH: 1389  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 PCT-US02-04915-41

## Alignment Scores:

Pred. No.: 4,66e-243 Length: 1389  
 Score: 2182.00 Matches: 407  
 Percent Similarity: 91.72% Conservative: 14  
 Best Local Similarity: 88.67% Mismatches: 24  
 Query Match: 89.28% Indels: 14  
 DB: 2 Gaps: 4

US-09-902-772-2 (1-451) x PCT-US02-04915-41 (1-1389)

QY 2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerIleuPheGluCys 21

|||  
 Db 25 GCAGCTCATATCAAGGAAGCCCTTATCACTTGTGATGAGACACAGTCTGTTTGTAGCT 84  
 QY 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41  
 |||  
 Db 85 GCCTACGGAAGCCACACCTGTGTAAAGACAGATGACGGGGTCTCTCTCCAGACATAT 144  
 QY 42 GlyAlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61  
 |||  
 Db 145 GGACGAGCTTCAGATAGGCCACGCGTCCCTCAGCAGGATGGCTCTTCACACCCCA 204  
 QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
 |||  
 Db 205 GCCAGGGTCACCATCAAAATGGAATGAACCTACCCAGGTGAATGGCTCAAGAACTCT 264  
 QY 82 ProAspArgCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101  
 |||  
 Db 265 CCTGATGAATACAGTGTGGCCAAAGCGGGAAGATGGTGGCAGCCAGACACCGTTGG 324  
 QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrAsn 121  
 |||  
 Db 325 ATGAACTACGCGCAGCTACATGAGAGAGACATGCAACCCCAACATACACAGAAC 384  
 QY 122 GluArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHisValArgGln 141  
 |||  
 Db 385 GAGGCAAGATTAATGTCGCCAGAGATCTCTAGCATGAGATACAGACCATGTGGCGAG 444  
 QY 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161  
 |||  
 Db 445 TGGCTGGAGTGGCGGTGAAGAATATGGCTTCAGACGTCACATCTGTATTCAG 504  
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181  
 |||  
 Db 505 AACATCGATGGGAAGAACTGTGCAAGTGAACAAAGGCGACTTCACAAAGCTCAACCCC 564  
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201  
 |||  
 Db 565 AGCTAACAAAGCGGACATCTCTCTCATATCTCCACATCTCCAGAGACATCCCTTCCA 624  
 QY 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrArgPro 221  
 |||  
 Db 625 CATTTG-----ACTTCA-----GATATGTGATAAAGCTTCAAAACTCTCA 669  
 QY 222 -----AspLeuProTyrGluGlnAlaIaArgArgSerAlaTrp 233  
 |||  
 Db 670 CGGTTAATGCTGCTAGAAACACAGATTTACCATATGAGCCGCCCAAGAGATCAGCTGG 729  
 QY 234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThrVal 252  
 |||  
 Db 730 ACCGGTACGGCCACGCCAGCCAGCCAGTCAAGTGTCAACCATCTCTCCACAGTG 789  
 QY 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272  
 |||  
 Db 790 CCAAAACTGAAGACCGGCTCTCAGTTAGATCTTATCAGATTCCTTGGACCAACAGT 849  
 QY 273 SerArgLeuAlaAsnProGluSerGlyGlnIleGlnLeuTrpGlnPheLeuGlnLeu 292  
 |||  
 Db 850 AGCCGCTTGCAATTCAGGCGAGTGGCCAGATCCAGCTTGGCACTTCTCTGAGCTC 909  
 QY 293 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlnIleTrpAsnGlyGlnPheLys 312  
 |||  
 Db 910 CTGTGGGACAGCTCACTCACTCAGCTGATCACTGGGAGGACCAACCGGGGATTTCAAG 969  
 QY 313 MetThrAspProaspGluValAlaIaArgArgTrpGluArgLysSerLysProAsnMet 332  
 |||  
 Db 970 ATGACGATCCCGACGAGGTGGCCGGCGGTGGGAGAGCGGGAAGACCAACCAACATG 1029  
 QY 333 AsnTyrAspLysLysSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352  
 |||  
 Db 1030 AACTACATTAAGCTCAGCGCGCTCGTTACTACTATGCAAGACATCATATCAACAG 1089  
 QY 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372  
 |||

```

Db      1090 GTCCATGGAGCGCTACGCCCTTACAGTTCACAGCGGATGCCCGCCAG 1149
Qy      373 ProHisProgluserSerMetTyrlYsTyProSerAspleuProTyMetSer 392
      1150 CCCCAGCCCCGGGATCTCTGTACAGAACCCCTCAGACCTCCGCTACATGGGCTCC 1209
Qy      393 TyHHisAlaHisProGlnLysMetAsnPhenValAlaProHisProProAlaLeuProVal 412
      1210 TATCACGCCCGACCCAGAGAGATGAACTTGTGGCGCCCGACCCCTCCAGCCCTCCCGT 1269
Db      413 ThSerSerSerPhePheAlaAlaProAsnProTyTrpAsnSerProThnglylyle 432
      1270 ACATCTTCCAGTTTGTGCTGCCCAACCCACTACTGGAATTCACCACTGGGGGTATA 1329
Qy      433 TyProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyTrpTytr 451
      1330 TACCCCAACACTAGGCTCCCGACACGCAATATGCTTCATCTGAGGCACTTACTAC 1386

RESULT 6
US-10-211-462-41
; Sequence 41, Application US/10211462
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Nafana
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-211-462-41

Alignment Scores:
Pred. No.: 4.66e-243 Length: 1389
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 89.67% Mismatches: 24
Query Match: 89.28% Indels: 14
DB: 14 Gaps: 4

US-09-902-772-2 (1-451) x US-10-211-462-41 (1-1389)
Qy      2 AlasSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerIleuPheGluCys 21
      25 GCAGCCTCATATCAAGAGAGCGCTTATCAGTTGTGAGTAGAGACCACTGCTTGTGATGT 84
Db      22 AlATyrlGlySerProHisIleuAlaLysThGluMetThrAlaSerSerSerGluTytr 41
      85 GCCTAGCGAAGCGCACACCTGCTGCTAAGACAGATGACCGCTCTCTCCACAGCACTAT 144
Qy      42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
      145 GGACAGAGCTTCCAGAGAGAGCCACGCGTCCCTCAGAGAGATTGGCTGTTCAACCCCA 204
Db      62 AlArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
      205 GCCAGGGTCCACCTCAAAATGGAATGATTAACCTTAGCCAGGTGATGCTCAAGGAACCTCT 264

```

```

Qy      82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
      265 CCGTATGAATGACAGTGTGGCCAAAGCGGGAAGATGTGGGACAGCCAGACACCGTTGGG 324
Db      102 MetAsnTyrlGlySerTyMetGluGluLysHisIleProProProAsnMetThrThAsn 121
      325 ATGAACACAGCGCACCTCATGAGAGAACACACATGGCACCCCAACATGACACAGAAC 384
Qy      122 GluArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHisValArgGln 141
      385 GAGCGCAGAGTTATGTCGCCAGCAGATCTTACGCTATGAGATGACAGACCATGTGGCCAG 444
Db      142 TrpLeuGlnTrpAlaValLysGluTyrlGlyLeuProAspValAspIleLeuLeuPheGln 161
      445 TGGCTGAGTGGGGGGTGAAGAATATGGCTTCACAGCTCAACATCTGTATTATTCAG 504
Qy      162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
      505 AACATCGATGGGAAGAGACTGTGCMAAGATGACCAAGAGACATTCACAGAGCTCACCCCC 564
Db      182 SerTyTrsnaAlaAspIleLeuLeuSerHisLeuHisIleTyTrpLeuArgGluArgLysAlaThr 201
      565 AGCTACAGACCGCCAGACATCTTCTCTCATCTCCACTCAGACAGAGACTCCTCTCCA 624
Qy      202 PheIlePheProAsnThrSerValTyTrpGluAlaThrGlnArgIleThrArgPro 221
      625 CATTTG-----ACTTCA-----GATGATGTTGATTAAGCCTTCAAAACTCTCCA 669
Db      222 -----AspleuProTyrlGlnGlnAlaArgAsnSerAlaTrp 233
      670 CGGTTAATGCATGCTAGAAACACAGATTTACCATATGAGCCCCCAGAGATATGACCTGG 729
Qy      234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThVal 252
      730 ACGGCTAGCGGCACACCCAGCCAGCCAGTGAAGAGCTCAACCATCTCTTCCACAGCTG 789
Db      253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrlGlnIleLeuGlyProThrSer 272
      790 CCCCAAACTGAAGACCAAGCTCTCTCACTGATGATTCATTATGATTCATGAGTTCAGCAACAAAGT 849
Qy      273 SerArgIleuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 292
      850 AGCGGCTTCAAAATCCAGGCAGAGGCGCAATCCAGCTTGGGAGTTCCTCGGAGCTC 909
Db      910 CTGTCGACAGCTCCCAACTCAGCTGATCAGCTGGGAGAGCCAGCAACAGGGGAGTTCAAG 969
Qy      293 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLysPheLys 312
      910 CTTGCGGACAGCTCCCAACTCAGCTGATCAGCTGGGAGAGCCAGCAACAGGGGAGTTCAAG 969
Db      313 MetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332
      970 ATGACGGATCCCGACGAGGTGGCGCGCTGGGGAGAGCGGAAGAACCAACCAACATG 1029
Qy      333 AsnTyTrsnaLysLysSerArgAlaLeuArgTyTrpTyTrsnaLysAsnIleMetThrLys 352
      1030 AACTAGCATTAAGCTCAGCGCGCTCTCGTTACTACTATGACAAAGAACATCATGACCAAG 1089
Db      353 ValHisGlyLysArgTyTrpAlaTyrlLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
      1090 GTCCATGGAGCGCTACGCCCTTACAGTTCACAGCGGATGCCCGCCAG 1149
Qy      373 ProHisProgluserSerMetTyrlYsTyProSerAspleuProTyMetSer 392
      1150 CCCCAGCCCCGGGATCTCTGTACAGAACCCCTCAGACCTCCGCTACATGGGCTCC 1209
Db      393 TyHHisAlaHisProGlnLysMetAsnPhenValAlaProHisProProAlaLeuProVal 412
      1210 TATCACGCCCGACCCAGAGAGATGAACTTGTGGCGCCCGACCCCTCCAGCCCTCCCGT 1269
Qy      413 ThSerSerSerPhePheAlaAlaProAsnProTyTrpAsnSerProThnglylyle 432
      1270 ACATCTTCCAGTTTGTGCTGCCCAACCCACTACTGGAATTCACCACTGGGGGTATA 1329

```

QY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451  
Db 1330 TACCCCAACACTAGGCTCCCAACGACCATATGCTTCTCATCTGGGACTTACTAC 1386

RESULT 7  
US-10-087-192-1127  
; Sequence 1127, Application US/10087192  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1127  
; LENGTH: 1389  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-087-192-1127

Alignment Scores:  
Pred. No.: 4,66e-243 Length: 1389  
Score: 2182.00 Matches: 407  
Percent Similarity: 91.72% Conservative: 14  
Best Local Similarity: 88.67% Mismatches: 24  
Query Match: 89.28% Indels: 14  
Gaps: 4

US-09-902-772-2 (1-451) x US-10-087-192-1127 (1-1389)

QY 2 AlaSerThrIleLeuGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21  
Db 25 GCAGCTCATATCAAGAAAGCCTTATCACTGTGAGTGGAGACCACTGTTGTAGCT 84

QY 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41  
Db 85 GCCTACGGAGAGCCACACCTGCTGATAGACAGATGACGGGCTCTCTCCAGGACAT 144

QY 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPheLeuSerGlnProPro 61  
Db 145 GGACAGACTTCCAGATGAGCCACGCGCTCCCTCAGCAGATTGGCTCTCAACCCCA 204

QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
Db 205 GCCAGGGTCCACATCAAAATGGAATGTAACCTTACGCCAGGTGAATGGCTCAAGAACTCT 264

QY 82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerSerAspAsnValGly 101  
Db 265 CCTGATGAATGCAGTGTGGCCAAAGCGGGAAGATGTTGGGGACAGCCAGACACCGTTGG 324

QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121  
Db 325 ATGACATACGGCAGCTACATGAGAGAGACATGACACCCCAAAACATACACACGAMC 384

QY 122 GluArgValIleIleValProAlaAspProThrLeuTyrPheSerHisValArgGln 141  
Db 385 GAGCGCAGAGTTATCTGTCACGACGATCTTACGCTATGAGATACAGACATGTGGGAG 444

QY 142 TrpLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161  
Db 445 TGGCTGAGTGGCGGGTGAAGAATATGCTTCACAGCGTCAACATCTTGTATTCAG 504

QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181  
Db 505 AACATCGATGGGAAGAACTGTGCAAGATGACCAAGGACGACTTCAGAGGCTCAACCC 564

QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgValThr 201  
Db 565 AGCTACAAACGCCGACATCTTCTTACATCTCCACTACCTACAGAGACATCTCTTCCA 624

QY 202 PheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThrArgPro 221  
Db 625 CATTTG-----ACTTCA-----GATGATGTGTGAATAAGCCTTCAAAACTCTCCA 669

QY 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTrp 233  
Db 670 CGGTAAATGATGCTAGAAACACACAGATTTACCATTTGAGCCCTCCAGAGATCACCTGG 729

QY 234 ThrSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThrVal 252  
Db 730 ACCGCTACGGCCACCCACGACCCGACGAGTGAAGCTGCTCAACATCTCTTCCACAGTG 789

QY 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272  
Db 790 CCCAAACTGAAGACACGCTCTCAGTTAGATCTTATTCAGATTCTTGGACCAACAGT 849

QY 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPheLeuLeuLeu 292  
Db 850 AGCCGCTTGCAAAATCCAGCAGTGGCCAGATCCAGTTTGCACTTCTCCTCGGAGCTC 909

QY 293 LeuSerAspSerSerAsnSerAsnCysTleThrTrpGluGlyThrAsnGlyGluPheLys 312  
Db 910 CTGTGGACAGCTCAACTCCAGCTGACATCACCTTGGAAGGACCAACGAGGAGTTCAG 969

QY 313 MetThrAspProAspGluValAlaLysArgTyrGlyGluArgLysSerLysProAsnMet 332  
Db 970 ATGACAGATGCCGACAGAGAGTGGCCCGGCGTGGGAGAGCGGAAGACCAACCAATG 1029

QY 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352  
Db 1030 AACTACGATTAAGCTCAGCGCGGCTCGCTTACTATGACAAACATCATCATGACCAAG 1089

QY 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372  
Db 1090 GTCCATGGGAAGCGTACGCTTACAGATTCGACTTCCACGGGATGCCAGGCTCTCAG 1149

QY 373 ProHisProGlnSerSerMetLysTyrTyrProSerAspLeuProTyrMetSerSer 392  
Db 1150 CCCACCCCGGAGATCATCTCTGTAAGAATACCCCTCAGACCTCCGTACATGGGCTCC 1209

QY 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412  
Db 1210 TATCAGCGCCACCCACACAAATGAATTTGTGGGCGCCACCCCTCCAGCCCTCCG 1269

QY 413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrGlyTyrIle 432  
Db 1270 ACATCTTCCAGTTTTTTTGTGCTGCCCCAACCATATCTGGAATTCACCACTGGGGTATTA 1329

QY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
Db 1330 TACCCCAACACTAGGCTCCCAACGACCATATGCTTCTCATCTGGGACTTACTAC 1386

RESULT 8  
US-10-126-052A-330  
; Sequence 330, Application US/10126052A  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Murray, Richard  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and  
; FILE REFERENCE: 018501-001530US  
; CURRENT APPLICATION NUMBER: US/10/126,052A  
; PRIOR FILING DATE: 2002-04-18  
; PRIOR APPLICATION NUMBER: US 60/284,770  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: US 60/290,492  
; PRIOR FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: US 60/339,245

```

; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,370
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 691
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 330
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-126-052A-330

Alignment Scores:
Pred. No.: 2,05e-242 Length: 3166
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 86.67% Mismatches: 24
Query Match: 89.28% Indels: 14
Gaps: 4

US-09-902-772-2 (1-451) x US-10-126-052A-330 (1-3166)
QY 2 AlaserthrlielysgluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
Db 281 GCACCTCATATCAAGAGAGCCCTTATCAGTGTGAGTGAGCAGCAGTCGTTGAGTGT 340
QY 22 AlatyrglySerProHisleuAlaValThrGluMetThrAlaSerSerSerGlyTyr 41
Db 341 GCCTACGAGACGCCACACCTGCTGCTAGACAGATGACGCGCTCTCTCCAGGACTAT 400
QY 42 GlylnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnProPro 61
Db 401 GGACAGACTTCCAGATGAGCCACGGCTCCCTCAGCAGATGGCTGTCTCAACCCCA 460
QY 62 AlAargValThrlielysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 461 GCCAGGGTCAACCAATGGAATGTAACTGACGAGGTGAATGGCTCAAGCAACTCT 520
QY 82 ProAspAPCysSerValAlaValGlyLysMetValSerSerSerAspAsnValGly 101
Db 521 CCTATGATGATGACATGTGGCCAAAGCGGGAAGATGGTGGCAGCCAGACCGGTGG 580
QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThrAsn 121
Db 581 ATGAACATGAGGACACTCATGTGAGAGAGACATGCAACCCCAACATGACACAGAC 640
QY 122 GluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArgGln 141
Db 641 GAGGACAGAGTTATCGCCAGCAGATCTTACGCTATGAGATGACAGACCATGTCCG 700
QY 142 TyrLeuGluTrrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db 701 TGGCTGAGATGGCGGGGGAAGATATGGCTTCCAGACGTCACATCTTGTATTCAG 760
QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 761 AACATGATGGGAAGGAAGATGTGCAAGATGACCAAGAGCATCTCCAGAGGCTCACCC 820
QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
Db 821 AGCTACACAGCCGACATCTTCTCTCACATCTCCACTACACTCAGAGAGACTCTCTCA 880
QY 202 PheIlePheProAsnThrSerValTyrProGlnLathrGlnArgIleThrThrArgPro 221
Db 881 CATTTG-----ACTTCA-----GATGATGTTGATTAAGCCTTACAAACTCTCA 925
QY 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTyr 233
Db 926 CGGTTATGATGCTAGAACACAGATTTACCATATGAGGCCGCCAGAGATTCAGCCTGG 985

```

```

QY 234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 252
Db 986 ACCGGTCACGGCCGCCACCCACCCAGTGGAAAGCTGCTCAACCATCTCTTCCACAGTG 1045
QY 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
Db 1046 CCCAAACTGAAAGCACAGCTGCTCAGTTAATCTTATGATTTCTTGACCAACAGT 1105
QY 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuGluLeu 292
Db 1106 AGCCGCTTGCAAATTCAGCAGCAGTGCAGATCCAGCTTGGACAGTTCCTCTGAGCTC 1165
QY 293 LeuSerAspSerSerAsnSerAsnGlyIleThrProGluGlnArgGlySerLysProAsnMet 312
Db 1166 CTGTGGACACTCCCACTCAGCTGATCCTGACCTGGGAAGCACCACAGGGGAGTTCAAG 1225
QY 313 MethrAspProAspGluValAlaArgArgTrrPglArgGluArgLysSerLysProAsnMet 332
Db 1226 ATGACGATCCCGACAGAGTGGCCGCTGGGGAGAGCGGAAGAGCAAAACCCACATG 1285
QY 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352
Db 1286 AACTACGATAGCTCAGCGCGCCCTCGTACTACTATGACAAAGAACATCATGACCAAG 1345
QY 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
Db 1346 GTCCATGGGAAGCCCTCAGCTTACAGTTCACCTTCCACAGGAGTCCGACGCCCTCCAG 1405
QY 373 ProHisProProGlnSerSerMetTyrLysTyrProSerAspLysProTyrMetSerSer 392
Db 1406 CCCACCCCGGAGTCACTCTCTACAGTACCCCTCAGCCTCCGTTATGAGCTCC 1465
QY 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412
Db 1466 TATCAGGCCACCCACAGAAAGATGACTTGTGGCGGCCACCCCTCAGGCCCTCCGCTG 1525
QY 413 ThrSerSerPhePheAlaAlaProAsnProTyrTrrPasnSerProThrGlyIle 432
Db 1526 ACATCTTCCACTTTTGTGCGCCCAAAACCATACGGAATTCACCAACTGGGGGTATA 1585
QY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db 1586 TACCCCAACTAGGCTCCACCAAGCATATGCTTCTCATCTGGGCACTTACTAC 1642

RESULT 9
US-10-087-192-1124
; Sequence 1124, Application us/10087192
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US 10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1124
; LENGTH: 1297
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-1124

Alignment Scores:
Pred. No.: 1.57e-228 Length: 1297
Score: 2057.00 Matches: 384
Percent Similarity: 83.69% Conservative: 11
Best Local Similarity: 81.36% Mismatches: 9
Query Match: 84.17% Indels: 68

```

DB: 15 Caps: 2  
 US-09-902-772-2 (1-451) x US-10-087-192-1124 (1-1297)

QY 7 GluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCysAlaTyrGlySerPro 26  
 |||||  
 Db 2 GAGGCGCTTGACAGTGTGACGAGGAGCAGTCACTATTGAGTGTGCTCAGGAAAGCCCA 61  
 |||||

QY 27 HisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlyGlyGlnThrSerLys 46  
 |||||  
 Db 62 CACCTGGCTTAAGACAGATGACCGCATCTCTCCAGTACTATGGCCAGACATCCAG 121  
 |||||

QY 47 MetSerProArgValProGlnGlnAspTyrLeuSerGlnProAlaArgValThrIle 66  
 |||||  
 Db 122 ATGAGTCCAGAGTCCCTCCAGAGAGTGGCTGTCTCAAGCCCGAGGAGGTCACCATC 181  
 |||||

QY 67 LysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAspCysSer 86  
 |||||  
 Db 182 AAGATGGAGTGAACCTTACTAGGTGAATGTTCAGGAAGTCCAGTGTGATGATGAGT 241  
 |||||

QY 87 ValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGlySer 106  
 |||||  
 Db 242 GTTAACAAAGGTGGGAAGATGTGGGCGACCGCATCTGTGGGATGAGCTACGGCAGC 301  
 |||||

QY 107 TyrMetGluGluTyrHisIleProProProAsnMetThrThrAsnGluArgValIle 126  
 |||||  
 Db 302 TACATGAGGAGGAAGCATGTGCGCCCTCCCAATATATACCAATATGAGCCAGAGTATC 361  
 |||||

QY 127 ValProAlaAspProThrLeuTyrPserThrAspHisValArgIleThrLeuGluTyrPala 146  
 |||||  
 Db 362 GTCCCTGCAGATCTCTCTGTGTGAGACAGACCATATGTCCGACAGTGGCGAGTGGCG 421  
 |||||

QY 147 ValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLys 166  
 |||||  
 Db 422 GTTAAGAAATATGCGCTCTCGATGTGAGCGTCTTACTATTGAGATATCGATGGGAG 481  
 |||||

QY 167 GluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAsp 186  
 |||||  
 Db 482 GAGCTGTGCAAGATGACAAAGATGACTTCCAGCGGTCTCAGCGGAGTCAATGCCGAC 541  
 |||||

QY 187 IleLeuLeuSerHisIleuHisTyrLeuArgGluArg 198  
 |||||  
 Db 542 ATTCTTCTCCACATCTCCACTACCTACAGAGACATCCCTCCACATCTGACTCCGAT 601  
 |||||

QY 199 199 Gly 199  
 |||||

Db 602 GACGTGATAGAGCTTTACAAACTCTCCAGGTTAATGATCCAGAAACAGAGGGGT 661  
 |||||

QY 200 AlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThr 219  
 |||||  
 Db 662 GCAGCTTTATTTTCCCAATACTTCAATATATCCGAAGCTTACGAAGATTTACACT 721  
 |||||

QY 220 ArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrThrSerHisSerHisPro 239  
 |||||  
 Db 722 AGGCCA 727  
 |||||

QY 240 ThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArg 259  
 |||||  
 Db 727 727  
 |||||

QY 260 ProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGly 279  
 |||||  
 Db 728 GATCCTTACCGATCCGATCCGAGCGAGCCAGTACCGCCTGTGTAATCCAGT 778  
 |||||

QY 280 SerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeuSerSerSerAsnSer 299  
 |||||  
 Db 779 AGTGGCCAGATCCAGCTGTGGAGTCTGCTCGAATCTCTGTCAGACAGACTCCCACTCC 838  
 |||||

QY 300 AsnCysIleThrThrPgluGluTyrThrAsnGlyGluPheLysMetThrThrAspProAspGluVal 319  
 |||||  
 Db 839 AACTGCAATCACTGGGGAAGCCACCAAGGGGAGTGAAGTGAACAACCCGAGGAGGTG 898  
 |||||

QY 320 AlaArgArgTyrPgluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArg 339  
 |||||

Db 899 GCTCGGCGCTGGGGGAGAGAGAGACACACCCCACTGAATATGACAAAGCTCAGCGCC 958  
 |||||

QY 340 AlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAla 359  
 |||||

Db 959 GCCCTCGCTTAATCTACGACAAACATCATGACCAAGGTGACAGGGAAGCCCTACGCC 1018  
 |||||

QY 360 TyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerSer 379  
 |||||

Db 1019 TACAAGTTTACTCTCCACAGGGAATGGCCAGGCGCTGAGCCACCTCTGAGTCGTC 1078  
 |||||

QY 380 MetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLys 399  
 |||||

Db 1079 CTGTACAGATACCCCTCCGACCTGCGCATATGAGGCTCTATACGCCACCCAGAG 1138  
 |||||

QY 400 MetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAla 419  
 |||||

Db 1139 ATGAACCTTGTGTCTCCACCTCCGCTCTCCAGTCAATCTTCCAGTCTTGTGT 1198  
 |||||

QY 420 AlaProAsnProTyrTyrPasnSerProThrGlyGlyIleTyrProAsnThrArgLeuPro 439  
 |||||

Db 1199 TCCCGGAACCATACTGTGAATTCACGACTGGGGGCAATCAACCGAAGACTAGGCTCCA 1258  
 |||||

QY 440 AlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451  
 |||||

Db 1259 GCCAGCATATGCCCTCTCTCTACCTGGGACCTACTAC 1294  
 |||||

RESULT 10  
 US-10-144-771-355  
 ; Sequence 355, Application US/10144771  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTUR, J. Craig  
 ; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF  
 ; FILE REFERENCE: CL001321  
 ; CURRENT FILING DATE: 2002-05-15  
 ; NUMBER OF SEQ ID NOS: 47235  
 ; SEQ ID NO 355  
 ; LENGTH: 2172  
 ; TYPE: DNA  
 ; ORGANISM: HUMAN  
 ; US-10-144-771-355

Alignment Scores:  
 Pred. No.: 1,99e-220 Length: 2172  
 Score: 1991.00 Matches: 374  
 Percent Similarity: 80.54% Conservative: 11  
 Best Local Similarity: 78.24% Mismatches: 9  
 Query Match: 81.46% Indels: 84  
 DB: 14 Gaps: 2

US-09-902-772-2 (1-451) x US-10-144-771-355 (1-2172)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 |||||

Db 107 ATGGCCAGCACATTAAGAGGCGCTGTGATGTGAGCGAGACAGCATCATTTGAG 166  
 |||||

QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 |||||

Db 167 TGTGCTTACGAGGAGCGCACACCTGGCTAAGACAGATGACCCGATCTTCCAGTAC 226  
 |||||

QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60  
 |||||

Db 227 TATGGCCAGACATCAAGATGAGTCCAGAGTCCCTCAGGAGACTGGCTGTCAAGCC 286  
 |||||

QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 |||||

Db 287 CCAGCCAGGCTCACATCAAGATGAGTGAACCTAGTAGAGTGAATGTTCCAGAGAC 346  
 |||||

QY 81 SerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnVal 100  
 |||||

Db 347 TCACCTGATGAGTGAAGTGTGAACAAAGGTGGGAAAGTGTGGCAGCCCGATACTGTTG 406  
 |||||



```

Db      532 GCCACGACCCCTTACAAACAGGAGTGCTGTGTACACACCTTACCTACGGGAA---588
Qy      199 GlnAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThr218
Db      589 ---AGTTCACTGCGGGCTTAATATACAACTCCACACCGCCACCAATCTCCACGATGAGT645
Qy      219 ThrArgProAsnLeuProTyrGlnGlnAlaArgSerIleAlaThrPheSerHisSerHis238
Db      646 GTCAAAAGAACCCCTTTTATGACTCAGTCAAGAGAGAGAGCTTGGGCAATACATGAAAT705
Qy      239 ProThrGlnSerIleValAlaThrGln---ProSerSerSerThrValProTyrThrGluAsp257
Db      706 TCTGGCCCTCAACAAAGAGCTCCCTTGAGGGGCGACAAACGATCAATAGAAATACAGAG765
Qy      258 GlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn277
Db      766 CAAGGGCCCGCCAGCAGATCCGTATCAGATCCTGGGCGCCGACACAGATCCGCTAGCCAC825
Qy      278 ProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeuLeuSerAspSerSer297
Db      826 CTTGAAAGCGGCGAGATCCAGCTGTGGCAATTCCTCTGGAGCTGCTCTCCGACAGCGCC885
Qy      298 AsnSerAsnGlyIleThrTyrGlnGlyThrAsnGlyIlePheLeuMetThrAspProAsp317
Db      886 AACGCCAGCTGTATCAGCTGGAGGAGGACCAACGGGAGTTCAAAATGACGGAGCCCGAT945
Qy      318 GlnValAlaArgAlaGlyTyrGlyGlnArgIleSerIleSerProAsnMetIleAsnTyrAspLeu337
Db      946 GAGGTGGCCAGGCGCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG1005
Qy      338 SerArgAlaLeuArgTyrTyrTyrAspIleValSerIleMetThrIleValHisGlyLeuArg357
Db      1006 AGCGGGCCCTCCCTTTTACTATGATATAAAACCTTTTACCAAAAGTCCAGCGGAGGAG1065
Qy      358 TyrAlaTyrIlePheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGln377
Db      1066 TATCTTACAAATTTGACTTCCAGCGATGCGATGCGAGCTGCGAGCCACATCCAGCGAG1125
Qy      378 SerSerMetTyrIleTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisPro397
Db      1126 TCGCCAGTACAAAGTACCTCTGACATCTCTTACATGCTTCTTACCAATGCCACAG1185
Qy      398 GlnIleMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhe417
Db      1186 CAGAGGTGAACCTTTGCTCCCTCCCATCCATCCCTCCATGCTGCTGCTCCATCCAGCTTC1245
Qy      418 PheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyIleTyrProAsn-----435
Db      1246 TTTGGAGCCGCAATCAATCAATGAGACTCCCGCCAGGGGGGAAATCTACCCCAACCCCAAC1305
Qy      436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleTyrTyrTyr451
Db      1306 GTCCCCCGCATCTTACACCCAGGCTGCTTACACTTAGGACACTTAC1356

RESULT 12
PCT-US02-41414-1209
: Sequence 1209, Application PC/TUS0241414
: GENERAL INFORMATION:
: APPLICANT: Morris, David
: APPLICANT: Engelhard, Eric
: TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
: FILE REFERENCE: A-71249/RMS/DCF
: CURRENT APPLICATION NUMBER: PCT/US02/41414
: CURRENT FILING DATE: 2002-12-26
: PRIOR APPLICATION NUMBER: US 09/747,377
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/798,586
: PRIOR FILING DATE: 2001-03-02
: NUMBER OF SEQ ID NOS: 1613
: SOFTWARE: Patentia version 3.1
: SEQ ID NO 1209
: LENGTH: 2957
: TYPE: DNA

```

```

: ORGANISM: Homo sapiens
PCT-US02-41414-1209

Alignment Scores:
Pred. No.: 1,97e-177 Length: 2957
Score: 1624.50 Matches: 306
Percent Similarity: 79.43% Conservative: 57
Best Local Similarity: 66.96% Mismatches: 83
Query Match: 66.47% Indels: 11
DB: 2 Gaps: 7

US-09-902-772-2 (1-451) x PCT-US02-41414-1209 (1-2957)

Qy      1 MetAlaSerThrIleIleValAlaLeuSerValValSerGluAspGlnSerLeuPheGlu20
Db      173 ATGAGCGGAGCTATTAAAGAGCTCTGTGGTGTGAGCGAGCAGCAGTCTCTTTTAC232
Qy      21 CysAlaTyrGly---SerProHisLeuAlaIleThrGlnMetThrAlaSerSerSer39
Db      233 TCAGCTTACGAGGAGCGGAGCCCATCTCCCAAGGCCGACATGATGCTCGGGAGTCT292
Qy      40 GlnTyrGlyGlnThrSerIleMetSerProArgValProGlnGlnAspTyrPLeuSerGln59
Db      293 GACTACGGGAGCGCCCAAGATCAACCCCTCCACACAGCAGAGTGGATCAATCAG352
Qy      60 ProProAlaArgValThrIleIleMetGluCysAsnProAsnGlnValAsnGlySerArg79
Db      353 ---CCAGTGAAGGCTCAACGTCACGCGGAGTAT---GACCAATGAATGATCCAGG403
Qy      80 AsnSerProAspAspCysSerValAlaIleGlyGlyIleValSerSerSerAspAsn99
Db      404 GAGTCTCGGAGTCTGAGCTCAGCGTATGCAATGCAAGCAGAGCTGGCGGAGCGAGTCC463
Qy      100 ValGlyMetAsnTyrGlySerTyrMetGluGluHis---IleProProAsnMet118
Db      464 AACCCATGAACATCAACAGCTATATGAGAGAGAGATGGCCCCCTCTCCCAACATG523
Qy      119 ThrThrAsnGlnArgValIleValProAlaAspProThrLeuTyrPserThrAspHis138
Db      524 ACCACCAACGAGAGAGATGATGCTCCCGCAGACCCACAGTGTGACACAGGACAT583
Qy      139 ValArgGlnTyrPLeuGlnTyrPAlaValIleGlyGlyIleLeuProAspValAspIleLeu158
Db      584 GTGAGGAGATGGCTGGAGTGGCCCATTAAGAGATGACAGCTGATGAGATGACACATCC643
Qy      159 LeuPheGlnAsnIleAspGlyIleGluLeuGluCysIleMetThrIleAspAspPheGlnArg178
Db      644 TTTTCCAGAAATGATGATGCAAGAGACTGTGTAATGAACAGAGAGAGACTTCTCCGC703
Qy      179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrIleArgGluArg198
Db      704 GCCACGACCCCTTACAAACAGGAGTGTGTGACACCTGATTAATCTTACCTAGGGAA---760
Qy      199 GlnAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThr218
Db      761 ---AGTTCACTGCTGGGCTTAATATACAACTCCACACCGCCACCAATCTCCATGATGAGT817
Qy      219 ThrArgProAsnLeuProTyrGlnGlnAlaArgSerIleAlaThrPheSerHisSerHis238
Db      818 GTCAAAAGAACCCCTTTTATGACTCAGTCAAGAGAGAGAGCTTGGGCAATACATGAAAT877
Qy      239 ProThrGlnSerIleValAlaThrGln---ProSerSerSerThrValProTyrThrGluAsp257
Db      878 TCTGGCCCTCAACAAAGTCTCCCTTGAGGGGCGACAAACGATCAATAGAAATACAGAG937
Qy      258 GlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn277
Db      938 CAAGGGCCCGCCAGCAGATCCGTATCAGATCCTGGGCGCCGACAGATCCGCTTAGCCAAC997
Qy      278 ProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeuLeuSerAspSerSer297
Db      998 CTGTGAGCGGCGGAGATCCAGCTGTGGCAATTCCTCTGAGAGTGTCTCCGACAGCGCC1057

```









```

Db      890  GTCAAGGAAGACCCCTTCTTATGACTCTGTGACGAGAGAGGATGAGCAATTAATATGAAC  949
QY      239  ProthrinInserlySAla---ThGlnProSerSerSerThrvAlProlystThGluasp  257
                :::::::::::::::  |||  :::::  |||  :::
Db      950  TCTGGCCTCAAAAGAGTCCTCCCTGGAGGATCAAGACCTAGGCAACACTGAG  1009
QY      258  GlAnrgrProGlnLeuAspProTygGlnIleleuGlyProThrSerSerArgLeuAlasn  277
                |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1010  CAGGGGCCCGCCAGCATCCTTATCAGATCTCTGGGGCCCAACCCAGCGCGCTAGCAAC  1065F
QY      278  ProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeuSerAspSerSer  297
                |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1070  CCTGGAGATGGGAGATCACTGAGCTGGGAGATTCTCTGGAAGTACTGTCCGACAGCGCC  1129F
QY      298  AsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLupelysmethAspProasp  317
Db      1130  AACCCAGCAGCTGTACACTCTGGAGGGGACCACAGCGGGAGTTCAAAATGACCGACCTGAT  1189F
QY      318  GluValAlaArGArGTTrpGlyLargLysSerLysProAsnMetAnTyAspLysLeu  337
                |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1190  GAGGTGGCCAGGCGCTGGGAGAGCGGAGAGCAAGCCCAACATGANTTATGACAACCTG  1249F
QY      338  SerArgAlaLeuArGTyTyTyTyAspLysAsnIleMetThrLysValHisGlyLysArg  357
Db      1250  AGCGGGGCGCCCTCCGATACTACTATGACAAAAACATTATGACCAAGTCATGGCAAAAG  1309F
QY      358  TyrAlaTyLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlu  377
                |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1310  TATGCTCAAGATTGACTTCATCCATGGCATTTGCCAGGCGCCGACGCCACATCCAAAGAG  1369F
QY      378  SerSerMetTyLysTyLysTyProSerAspLeuProTyLysMetSerSerTyHisAlaHisPro  397
                :::::::::::::::  |||  :::::  |||||  |||||  |||||  |||||  |||||
Db      1370  ACATCCATGTAACAAGTATCCCTGTGAATATCCCAATGATGCCCTCTCCACAGGCCATCAA  1429F
QY      398  GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhe  417
                |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1430  CAGAAAGGTGAACCTTGTGCCGCTCCACCCACATCCATGCCGTGTACACTCCTCCAGCTTC  1489F
QY      418  PheAlaAlaProAsnProTyLysTrpAsnSerProThrGlyLysIleTyProAsn-----  435
                |||  |||  ::  |||||  |||||  |||||  |||||  |||||  |||||
Db      1490  TTGTGGACAGATCCACATACTAGTGAACCTCCGCCCACTCGGAGATCTATCCAAACCCAGT  1549F
QY      436  ---ThrvArgLeuProAlaAlaHisMetProSerHisLeuGlnIlyThrTyTyTy  451
                |||  |||  |||  |||  ::  ::::  ::::  ::::  ::::  ::::  ::::
Db      1550  GTCCCGCGCATCTTAACACCCAGCTGCTTCAACACTTAGGACACTACTAC  1600

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

3853.895 Million cell updates/sec

Sequence: 1 MASTIKKALSVSEDSLE.....IYPNTRLPAAHMPHGLTYY 451

Xgapop	10.0	,	Xgapext	0.5
Ygapop	10.0	,	Ygapext	0.5
Fgapop	6.0	,	Fgapext	7.0
Delop	6.0	,	Delext	7.0

Total number of hits satisfying chosen parameters: 32308132

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

```
-MODEL=trimeat_p2n.model -DEV=xlh
-o/cqrg2.1/USPTO.spool/USO9902772/r/unat.23072003.036567.14892/app.query.fasta_1.1299
-DE=EST -OFMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blotsum62 -TRANS=human0.0.cdi -LIST=45
-DUALALIGN=200 -THE_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=mc -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=USO9902772.@CGC 1.1.2619 @runat.23072003.036567.14892 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBIO=100 -LONGIOS
-DEV.TIMEDOUT=120 -WARN.TIMEDOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELXT=7
```

Database :

```

1: em_estbo:*
2: em_estbm:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_estl:*
10: gb_estr:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_hun:*
20: em_gss_pl:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
---------------	-------	----------------	--------	----	----	-------------

1	1118.5	45.8	827	9	AJ456498	AJ456498	AJ456498
---	--------	------	-----	---	----------	----------	----------

3	1016.5	41.6	880	14	B0233264	B0233264	AGENCOUT
4	1002.5	41.0	1089	14	B0212127	B0212127	AGENCOUT

5	999	40.9	715	12	BG388025	BG388025	60241280
---	-----	------	-----	----	----------	----------	----------

7	984	40.3	642	13	BI713036	BI713036	1e01d08
---	-----	------	-----	----	----------	----------	---------

8	984	40.3	668	10	BB660034	BB660034	BB660034
9	980	40.1	592	13	BM088063	501436	BM088063
10	975.5	39.9	915	12	BG390291	602416291	BG390291
11	971.5	39.8	701	9	AU136709	AU136709	AU136709

13	966	39.5	1074	14	BQ953977	BQ953977	AGENCOUR
13							

15	947.5	38.8	634	13	BM489636	BM489636	pgm2n.p)
15	947.5	38.8	634	13	BM489636	BM489636	pgm2n.p)

17	932	38.1	629	10	BB577887	BB577887
18	840.5	34.4	671	13	BI558530	BI558530 6032405

20	822	33.6	553	12	BG256864	BG256864	60237111
----	-----	------	-----	----	----------	----------	----------

22	816.5	33.4	524	9	AI794504	AI794504	fc45b11.y
22	816.5	33.4	524	9	AI794504	AI794504	fc45b11.y

24	793	32.4	522	12	BF566705	BF566705	UI-R-BJC
----	-----	------	-----	----	----------	----------	----------

[illegible]

36	620	25.4	894	14	BQ439561	BQ439561	AGENCOUR
----	-----	------	-----	----	----------	----------	----------

37	616.5	25.2	756	9	A1A45352	A1A45352
38	616.5	25.2	815	9	A1A46143	A1A46143
39	611	25.1	80755477	14	B0715477	NCENCOU
40	611	25.0	465	14	R87572	R87572
41	610.5	25.0	967	14	BQ642521	Ym8906
42	507.5	24.9	787	10	BE539913	AGENCOUT
43	506.5	24.9	587	9	AL539819	601345463
44	595.5	24.4	551	9	AA467733	DKFZP1313
45	585.5	24.0	788	9	AJ456869	V60206
						AJ456869

## ALIGNMENTS

RESULT 1	
LOCUS	AI456498
DEFINITION	827 bp mRNA linear EST 22-APR-2002
ACCESSION	AI456498
VERSION	AI456498.1
KEYWORDS	GI:20266594
SOURCE	EST.
ORGANISM	chicken.
	Gallus gallus

REFERENCE  
AUTHORS  
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 827)  
Buerstede, J.M.  
Gallus gallus bursal lymphocyte EST

JOURNAL Unpublished (2002)  
 COMMENT Contact: Buerstedde JM  
 Cellular Immunology  
 Heinrich-Pette-Institute  
 Martinistr. 52, 20251 Hamburg, Germany  
 Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.  
 Location/Qualifiers

FEATURES  
 source  
 1..827  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"  
 /clone="Bm13r2"  
 /cell\_type="bursal lymphocyte"  
 /dev\_stage="2-3 weeks old"  
 /note="CB inbred strain"

BASE COUNT 233 a 236 c 177 g 181 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2,21e-103 Length: 827  
 Score: 1118.50 Matches: 211  
 Percent Similarity: 78.31% Conservative: 20  
 Best Local Similarity: 71.53% Mismatches: 43  
 Query Match: 45.77% Indels: 21  
 DB: 9 Gaps: 3

US-09-902-772-2 (1-451) x AU456498 (1-827)

OY 138 HisValaArggIntPleuGluTPalaValaLysGluTyrglyLeuProAspValaSpile 157  
 |||||  
 DB 5 CACGTGCCCATGTGGCTGGATGGCCATTAAGGAGTATGATTAATGACATGACACC 64  
 |||||  
 OY 158 LeuPhegInaSnileAspGlyLysGluLeuCySlySmetThrlyAspAspPheGln 177  
 |||||  
 DB 65 ACCCTCTCCAGAAATGATGATGCAAGAGCTCTGCAGAAAGACAAAGATGACTCTCTC 124  
 |||||  
 OY 178 ArgLeuThrProSerTyraSnAlaAspIleLeuSerHisLeuHisTyLeuArgGlu 197  
 |||||  
 DB 125 CGAACCACTCCCTCTACACACAGAACTTGTGTCTCACCTCAGTTACCTCAGGGAA 184  
 |||||  
 OY 198 ArgGlyAlaThrPheIlePheProAsnThrSerValTyrrProGluAlaThrGlnArgIle 217  
 |||||  
 DB 185 AGTAGCTACGTGCTGTAC---AATTACTCATCCACACAGAGACTCTCTCAGTCTT 241  
 |||||  
 OY 218 ThrThrArgProAspLeuProTyrgluGlnAlaArgSerAlaTrpThrSerHisSer 237  
 |||||  
 DB 242 GCCACCAAGAAAGGTCTCTCTGT----- 265  
 |||||  
 OY 238 HisProthrGlnSerLysAlaThrGlnProSerSerSerThrValProlyThrGluAsp 257  
 |||||  
 DB 266 -----GCAGGGACACAAACGCTGAACAGACA-----ACAGAACAG 301  
 |||||  
 OY 258 GlnArgProGlnLeuAspProTyrglnIleLeuGlyProThrSerSerArgLeuAlaSn 277  
 |||||  
 DB 302 CAAGGGCTCAGCAGATCCCTATCAAAATCTGGGGCCACTAGTAGTGTCTTGCAT 361  
 |||||  
 OY 278 ProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSer 297  
 |||||  
 DB 362 CCTGGAGTGGGAGATACAACTATGGAGTTCCTCCGAGTTCGTCGAGACAGTTC 421  
 |||||  
 OY 298 AsnSerAsnCySlyIleThrTrpGlyGlyThrAsnGlyLeuPheLySmetThrAspProAsp 317  
 |||||  
 DB 422 AACCCACGCTGTATCATATGGAGGACCAATGGGGAATTCAGAGATGACAGCCAGAT 481  
 |||||  
 OY 318 GlnValAlaArgArgTyrglyGlnArgLysSerLysProAsnMetAsnTyraSpLysLeu 337  
 |||||  
 DB 482 GAATGGACGCGCTGGGGAGAACGCAAGACCAAGCCCAATGATTAATGACAGAGTGG 541  
 |||||  
 OY 338 SerArgAlaLeuArgTyrglyTyraSpLysAsnIleMetThrlyValHisGlyLysArg 357  
 |||||  
 DB 542 AGCGAGCCCTTCATCTACTATGATGAAGACATTTAGCAAGATGCAAGGCAAGCA 601  
 |||||  
 OY 358 TyraAlaTyrllysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGln 377  
 |||||

DB 602 TATGCCCTACAAATTGACTTTCATGCGATTCGCCAGGCTCTCCACCTTCATCCCACTAA 661  
 |||||  
 OY 378 SerSerMetTyrlsTyrrProSerAspLeuProTyrrMetSerSerTyrrHisAlaHisPro 397  
 |||||  
 DB 662 TCATCATATGTCACAGATATCCATCAGATCTCTCTCTCATGATCCCTTTCACATGCCACAG 721  
 |||||  
 OY 398 GlnLysMetAsnPheValAlaProHisProAlaLeuProValThrSerSerPhe 417  
 |||||  
 DB 722 CAGAAAGTGAAGCTTGTATCCGCCACACCTCTCTTATAGCCGTGCATCATATCAGTTTC 781  
 |||||  
 OY 418 PheAlaAlaProAsnProTyrrTrpAsnSerProThrGlyGlyIle 432  
 |||||  
 DB 782 TTCGGAGCAGGCTCACCCTATTTGACCTCTCTGTCTGGAGACATT 826  
 |||||

RESULT 2  
 BM456833  
 LOCUS  
 DEFINITION  
 5', mRNA sequence.  
 ACCESSION  
 BM456833  
 VERSION  
 BM456833.1 GI:18505873  
 KEYWORDS  
 EST.  
 SOURCE  
 ORGANISM  
 human.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC http://mgc.nci.nih.gov/.  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: c9apbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
<http://image.llnl.gov>  
 Plate: L14M12347 row: f column: 10  
 High quality sequence stop: 669.  
 Location/Qualifiers

FEATURES  
 source  
 1..1045  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5583753"  
 /clone\_lib="NIH-MGC-92"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: testis; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 2.5 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH-MGC Library."

BASE COUNT 300 a 290 c 230 g 225 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3.13e-95 Length: 1045  
 Score: 1040.00 Matches: 214  
 Percent Similarity: 75.21% Conservative: 17  
 Best Local Similarity: 69.74% Mismatches: 33  
 Query Match: 42.55% Indels: 44  
 DB: 13 Gaps: 8

US-09-902-772-2 (1-451) x BM456833 (1-1045)

OY 7 GlnAlaLeuSerValValSerGluAspGlnSerLeuPheGluCysAlaTyrglySerPro 26  
 |||||  
 DB 190 GAACCTTATACAGTTGGAGTGAAGACACAGTCGTTTGTAGTGTGCTCAGCAACGCCA 249  
 |||||  
 OY 27 HisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlyTyrglyGlnThrSerLys 46  
 |||||

```

Db 250 CACCTGGCTAGACAGAGATACCGCGCTCCTCCACGACCTATMGACACACTTCCAG 309
Qy 47 MetSerProArGValProGInGlnAspTrLeuSerGlnProProAlaArGValThrLe 66
Db 310 ATGAGCCACACCGCTCCACAGAGATTGGCTGTCTCAACCCACGAGGTCACACATC 369
Qy 67 LysMetGluCysAsnProAsnGlnValAsnGlySerAspAsnSerProAspCysSer 86
Db 370 AAATGGATTAACCTACGACAGTGAATGCTTCAGAGAACTCTCCGATGAATGCAGT 429
Qy 87 ValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTrGlySer 106
Db 430 GTGGCCAAAGCGGGAGAAATGCTGGCGACCCACACACCGTTGGATGAATACGCGCAGC 489
Qy 107 TyrMetGluGluLysHisIleProProProAsnMetThrThrAsnGluArGValIle 126
Db 490 TACATGGAGGAGAGACATGACACCCCAACATGACACGAGAACGAGCGCATATC 549
Qy 127 ValProAlaAspProThrLeuTrpSerThrAspHisValArGValTrpLeuGluTyrPala 146
Db 550 GTGGCAGACATCTTACGCTATGAGATACACATGTCGCGCAGTGGCTGAGTGGCGC 609
Qy 147 ValLysGluTrpGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLys 166
Db 610 GTGAAGAATATGGCTTCCAGACGTCACATCTTGTATTCAGAGAACATGATGGGAG 669
Qy 167 GluLeuCysLysMetThrLysAspAspPheGlnArGValTrpProSerTyrAsnAlaAsp 186
Db 670 GAATGTGACAGATGACCCAGACACTCCAGAGGCTCACCCCGCTACACGCGCAC 729
Qy 187 IleLeuLeuSerHisLysHisTyrLeuArGValArGValArGValPhePheProAsn 206
Db 730 ATCTTCTCTACATCTTCACTACCTCAGAGACGCTCTTCCACATTTACCTCCAGAT 789
Qy 207 ThrSerValTyrProGlnAlaThrGlnArGValIleThrThrArPro----- 221
Db 790 -----GATGTGTGATAAAGC-CTTACAAACCTCCCGGTATATCATGCT 833
Qy 222 -----AspLeuProTyrGluGlnAlaAlaArGValArGValTrpHisSerHis 238
Db 834 AGAAACACCGATTTACATATGAGACCCCGGAGATCAGCTGACCGGTCAACGCGTCA 893
Qy 239 ProThrGlnSerLysAlaThrGlnProSerSerSerThrVal-----Pro 253
Db 894 CCC-----CAAGCCCAAGTGAAGAGCTTTCAACATCTTCCCTTCAAGGGGCC 944
Qy 254 LysThrGluAspGlnArGValProGlnLeuAspProTyrGlnIleGluGlyProThrSerSer 273
Db 945 AAAACTGAA-----AGACCCAGC-----GGCCCTCAGGT--- 974
Qy 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeuLeu 293
Db 975 -----TAAATTCCT-----TTAATCAATCATCTT 998
Qy 294 SerAspSerSerAsnSerAsn 300
Db 999 TGGGACACCCACATATTAAC 1019

```

```

RESULT 3
BQ233264 880 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT 7565836 NIH_MGC_92 Homo sapiens cDNA IMAGE:6041526
DEFINITION 5', mRNA sequence.
ACCESSION BQ233264
VERSION BQ233264.1 GI:20414664
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 880)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

```

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMU at:
http://image.lnl.gov
Plate: LAM13279 row: h column: 07
High quality sequence stop: 608.
Location/Qualifiers
1. 880
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6041526"
/clone_id="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NOTI;
Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 250 a 231 c 207 g 189 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 5.98e-93 Length: 880
Score: 1016.50 Matches: 195
Percent Similarity: 84.94% Conservative: 8
Best Local Similarity: 81.59% Mismatches: 7
Query Match: 41.59% Indels: 29
Db: 14 Gaps: 2
US-09-902-772-2 (1-451) x BQ233264 (1-880)
Qy 1 MetLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 74 ATGGCAGACATATTAAGGAGCCTTATCAAGTTGTGAGTGAGACAGTGGCTGTTCAG 133
Qy 21 CysAlaTyrGlySerProHisLysLeuAlaLysThrGlnMetThrAlaSerSerSerGlu 40
Db 134 TGTGCTACGGAAGCCACACCTGCTAAGACAGATGACCGGCTCTCTCCAGGAC 193
Qy 41 TyrGlyLysThrSerLysMetSerProArGValProGlnGlnAspTrpLeuSerGlnPro 60
Db 194 TATGACAGACTTCCAGATGAGCCACGCGCTCCAGCAGATGGCTGTCAACCC 253
Qy 61 ProAlaArGValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerAsn 80
Db 254 CCAGCCAGGGTCACACATCAAAATGGAATGTAACCTTGCAGCGGGAATGGCTCAAGAAC 313
Qy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db 314 TTTCTGATGATGACAGTGTGGCCAAAGCGGGAGATGTGTGGCAGACCCACAGACCGCT 373
Qy 101 GlyMetAsnTrpGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
Db 374 GGGATGAACATAGCGCAGCTCATGAGAGAGACATGACACCCCAACATGATGACAGC 433
Qy 121 AsnGluArGValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
Db 434 AACGAGCGCAGAGATTATCTGCCAGCAGATCTTACCTATGAGATTACAGACATGTGCGG 493
Qy 141 GlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuPhe 160
Db 494 CAGTGGCTGAGATGGGGGTGAAGATATGGCTTCCACACGTCACATCTTGTATATC 553
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArGValThr 180

```

Db 554 CAGAACATCGATGGGAGGAAGACTGTGCAGATGACCAAGACGACTTCCAGAGGCTCACC 613  
 QY 181 ProserThyrAsnAlaAspIleLeuSerHisLeuHisTyrLeuArgGluArg----- 198  
 Db 614 CCCAGCTACAGCGCCGACATCTTCTCTCACATCTCCACTACAGAGAGACTCTT 673  
 QY 198 ----- 198  
 Db 674 CCACATTTGACTTCAGATGATGNTGATAAACCCTTACAAACTCTCCACGGGNTAATGC 733  
 QY 199 -----GlyAlaThrPheIlePheProAsn---ThrSerValTyr 210  
 Db 734 ATGGCTAGAAACACAGAGGGGGTGCAGCTTTTATTATTTCCCAATTAATCTCAGTATAT 790  
 RESULT 4  
 B0212127 1089 bp mRNA linear EST 02-MAY-2002  
 LOCUS AGENCOURT\_7571287 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:6052136  
 DEFINITION B0212127  
 ACCESSION B0212127  
 VERSION B0212127.1 GI:20392058  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 1089)  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rgabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LLM13307 row: b column: 09  
 High quality sequence stop: 562.  
 Location/Qualifiers  
 1..1089  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6052136"  
 /clone\_lib="NIH\_MGC\_92"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: testis; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; Oligo-dT primed.  
 Average insert size 2.5 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 292 a 314 c 257 g 225 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.2e-91 Length: 1089  
 Score: 1002.50 Matches: 207  
 Percent Similarity: 70.31% Conservative: 18  
 Best Local Similarity: 64.69% Mismatches: 45  
 Query Match: 41.02% Indels: 50  
 DB: 14 Gaps: 5  
 US-09-902-772-2 (1-451) x B0212127 (1-1089)  
 QY 1 MetaLaserThrllelYsguAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 76 ATGGCCAGCACTATTAAAGAGCCTTATCAGTGTGAGTGAGCAGCAGTCCTTTGAG 135  
 QY 21 CysAlaIyrGlySerProHisLeuAlaIysThrGluMetThrAlaSerSerSerGlu 40

Db 136 TGTCCTACAGGAGCCACACACTGGCTTGAAGACAGAGATGACCGGCTCTCTCCAGGAC 195  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 60  
 Db 196 TATGACAGACTTCCAAAGATGAGCCACAGCGTCCCTCANCAGATGTGGTGTCTCAACCC 255  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 256 CCAGCCAGGGCTCACCATCAAAATGAAATGTAACCTAGCCAGGTAATGCTCAAGGAC 315  
 QY 81 SerProAspArgCysSerSerValAlaLysGlyGlyLysMetValSerSerSerAspAspVal 100  
 Db 316 TCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProPropoAsnMetThrGly 120  
 Db 376 GGGATGAACCTACGCGACCTACATGAGAGAGAGACATGCCACCCCAACATACACAGC 435  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140  
 Db 436 AACGAGCGCAGAGATTATCTGCCAGAGATCTTACGCTATGAGTACAGACCATGTGCGG 495  
 QY 141 GlnThrLeuGluTrrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160  
 Db 496 CAGTGGCTGGAGTGGGGGCTAAAGAAATATGCTTCCAGACGTCAACATCTTGTATTC 555  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 556 CAGAACATCCATGGGAAGAACTGTGCAAGATGACCAAGCAGACTTCCAGAGGCTCACC 615  
 QY 181 -ProserThyrAsnAlaAspIleLeuSerHisLeuHisTyrLeuArgGluArg----- 198  
 Db 616 CCCAGCTACATGGCCGACATCTCTCTCATCTCCTCAGAGAGACTCTCT 675  
 QY 198 ----- 198  
 Db 676 TCACATTTGACTACATGATGCTGTGATAAGCTTTCAAACTCTCCACGGGTAATGC 735  
 QY 199 -----GlyAlaThrPheIlePheProAsnThr-----SerValTyr 210  
 Db 736 ATGGCTAAACCCAGGGGGGCGCAGCTTTTATTTTCCCAATTCATTCAGAAATATTC 795  
 QY 210 rProGluAlaThrGlnArgIleThrThrArg-PropAspLeuProTyrGluGlnAlaArg 230  
 Db 796 GGAAGATCAGCCCAAAATTTACACCTAGGCGCAAAATTTACCTTATGAGCCCCCCC 855  
 QY 230 rg-----SerAlaThrThr-SerHisSerHisProThrGlnSerLysAlaThrGlnPro 247  
 Db 856 GGAAGATCAGCCCTGGACCGGTCATGGCCACCC----- 891  
 QY 248 SerSerSerThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIle 267  
 Db 892 -----CCGAGCGCCGAGGGTGAAGAAAGCGTGGCCCCCAATTTCCCCCT----- 933  
 QY 268 LeuGlyProThrSerSerArgLysLeuAlaAsnProGlySerGlyGlnIleGlnLeu 285  
 Db 934 ---TCCCGGGGGGCCCAAACTGAAGAACACAGCGGTGCCCTTGTTAAATC 984  
 RESULT 5  
 Bg388025 715 bp mRNA linear EST 12-MAR-2001  
 LOCUS 602412867F1 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:452123 5'  
 DEFINITION Bg388025  
 ACCESSION Bg388025  
 VERSION Bg388025.1 GI:13281471  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 715)  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: c9abs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM10420 row: f column: 18  
High quality sequence stop: 693.  
Location/Qualifiers

FEATURES  
source

1. 715

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4521233"  
/clone\_id="NIH\_MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: PCMV-SPORT6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC library."

BASE COUNT 206 a 188 c 172 g 149 t  
ORIGIN

Alignment Scores:

Pred. No.: 2.64e-91 Length: 715  
Score: 999.00 Matches: 186  
Percent Similarity: 97.46% Conservative: 6  
Best Local Similarity: 94.42% Mismatches: 5  
Query Match: 40.88% Indels: 0  
Gaps: 12 0

US-09-902-772-2 (1-451) x BG388025 (1-715)

QY 1 MetlaserthriletyglualaleuserValValsergluaapginserleupheglu 20  
Db 94 ATGCCACGACCTATTAGGAAGCCTTATCACTTGTGATGAGACCAAGCTTTTGAG 153  
QY 21 CysAlaTyrglySerProHISleuAlaIysThrglunetThralaserSerSerSerglu 40  
Db 154 TGTGCTACGGAAGCAGACACCTGTGCTAAGACAGATGACCGGCTCTCTCCAGGAC 213  
QY 41 TyrglylnhrserlysmetSerProArValProgluInasprtleuSerGlnpro 60  
Db 214 TATGACAGACTTCCAGATGAGCCACGGCTCCCTCAGCAGATGGCTGTCAACCC 273  
QY 61 ProAlaArValThriletysmetglucysasnproasnnglnvalasnnglyserArasn 80  
Db 274 CCAAGCCAGGGTCACCATCAAAATGAAATGATACCTTAGCCAGATGAATGGCTCAAGGAC 333  
QY 81 SerProaspapcysserValAlaIaIyglYlysmetValSerSerSerAspAsnVal 100  
Db 334 TCTCCTGATGAATGCAATGTGGCCAAAGGGGGAAGATGTGGGACCCAGACACCGTT 393  
QY 101 GlymetAsnTyrglyserTyrmetyglululysHisleProProProAsnmetThr 120  
Db 394 GGGATGACACGCGCAGCTACATGAGAGAGACATGCCACCCCAACATGACACAGC 453  
QY 121 AsnGluArArValIleValProAlaAspProthrleuThrserThrAspHisValArg 140  
Db 454 AACGAGGCGAGAGTTATGTCAGCAGATCTACGCTATGAGATAGACCAAGCTGCGG 513  
QY 141 GlnThrleuGlnTrpAlaValIyglulYtyrglyleuProAspValAspIleleuLeuPhe 160  
Db 514 CAGTGGCTGGAGGGGGGGAAGATATGGCTTCACAGCTCAACATCTTATTC 573  
QY 161 GlnAsnIleAspGlylysgluLeuCysLysmetThrLysAspAspPheGlnArgleuThr 180  
Db 574 CAGAACATCATGTGGAAGAACTGTGCAAGATGACCAAGAGACATCTCCAGAGGCTCAAC 633

QY 181 ProserTyraAlaAspIleleuLeuSerHisIstYrleuArggluArglu 197  
Db 634 CCCAGCTACAAAGCCGACATCTCTCTCAATCTCCATATCCATACAGAG 684

RESULT 6  
LOCUS B1558012  
DEFINITION 949 bp mRNA linear EST 05-SEP-2001  
603240720F1 NCI\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:5293492 5',  
mRNA sequence.

ACCESSION B1558012  
VERSION B1558012.1 GI:15445326  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
EMAIL: c9abs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM11742 row: h column: 05  
High quality sequence stop: 876.  
Location/Qualifiers

1. 949  
/organism="Mus musculus"  
/strain="NMRI"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5293492"  
/clone\_id="NCI\_CGAP\_Mam4"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: PCMV-SPORT6; Site: 1; SalI; Site: 2; NotI; Cloned unidirectionally. Primer: oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

BASE COUNT 252 a 288 c 209 g 200 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.08e-89 Length: 949  
Score: 985.00 Matches: 190  
Percent Similarity: 79.71% Conservative: 30  
Best Local Similarity: 68.84% Mismatches: 47  
Query Match: 40.30% Indels: 9  
Gaps: 3

US-09-902-772-2 (1-451) x B1558012 (1-949)

QY 183 TyraAsnAlaAspIleleuLeuSerHisIstYrleuArggluArgluPhe 202  
Db 2 TACAAACACAAAGCTGTTGCGACATCAGTTACCTCAGGAA-----ACTTACTG 55  
QY 203 IlePheProAsnThrSerValTyrrProglAlaThrGlnArgrIleThrThrArgProasp 222  
Db 56 CTGGCCATATACACACACCTCCATACAGACACAGTCTCCACAGCTGAATGCAAGAGAC 115  
QY 223 LeuProTyrglnGlnAlaArGargSerAlaTrpThrSerHisSerHisProthrGlnSer 242  
Db 116 CTTTATATGACCTGTGACGAGAGAGCATGMAACAATATATGAACTGTGGCCTCAAC 175





Db	532	CAGAAACATCATTGGGAGGACACTGTGCAGAATGACCAGACGACTTCACCAAGCTCAC	591	
Oy	181	ProSeTyRNA[aa]aspIleLeuLeuSerHisLeuHisTyrLeuArgLgU	197	
		::::::::::::::::::::		
Dd	592	CCCGAGTCANACGCCGACATCCTTCTCTCATCATCTCCACATCCACAGGA	642	
RESULT 8				
LOCUS	B8660034			
DEFINITION	B8660034 RIKEN full-length enriched, 13 days embryo lung Mus			
ACCESSION	B8660034			
VERSION	B8660034			
KEYWORDS	B8660034.1 GI:16493855			
SOURCE	EST.			
ORGANISM	house mouse. Mus musculus			
REFERENCE	Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Himoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda, M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasakiki, D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.			
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, URI:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kirita.A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1753-1771 (2000) Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara Y., and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library.Genome Res. 11 (2), 281-289 (2001) Konodo,S., Shingawa,A., Saito,T., Kiysawa,H., Yamana,I., Aizawa K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata.K. and Hayashizaki,Y. Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site ( <a href="http://genome.gsc.riken.go.jp">http://genome.gsc.riken.go.jp</a> ) for further details. e mouse tissues. Location/Qualifiers 1..668 /organism="Mus musculus" /db_xref="taxon:10090" /clone_id="D430027L04" /clone_lib="RIKEN full-length enriched, 13 days embryo lung" /tissue_type="Lung" /dev_stage="13 days embryo" /lab_host="DH10B" /note="Site_1: SalI; Site_2: BamHI; cDNA library was			

[illegible]

KEYWORDS EST.  
SOURCE COW.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 592)  
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett  
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,  
Perlee,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and  
Keel,J.W.  
Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
21180013  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt. trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.  
PCR Primers  
FORWARD: AGGAACACGCTATGACCAT  
BACKWARD: GTTTCCTCAGTCACGACG  
Plate: 139 row: D column: 21  
Seq primer: ATTTAGGTGACACTATAG.  
FEATURES  
source Location/Qualifiers  
1..592  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_1lb="MARC 2BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: PCMV SPORT6; Site.1: NotI; Site.2: SalI;  
Library made from pooled tissue from testis, thymus,  
semitendinosus muscle, longissimus muscle, pancreas,  
adrenal, and endometrium."  
BASE COUNT 152 a 177 c 159 g 104 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,71e-89 Length: 592  
Score: 980.00 Matches: 181  
Percent Similarity: 96.89% Conservative: 6  
Best Local Similarity: 93.78% Mismatches: 6  
Query Match: 40.10% Indels: 0  
DB: 13 Gaps: 0  
US-09-902-772-2 (1-451) x BM088063 (1-592)  
QY 1 MetAlserThrIleYsgIAlaIeuserValSerGlnuSpGInserleuPheGlu 20  
Db 12 ATGGCCAGCAGCTATTAAAGGCGTATCATCGTGTGAGCAGACCACTGCTTTGAG 71  
QY 21 CysAlaTyrGlySerProHisIleuAlaTyrThrGlnuThrAlaSerSerSerGlu 40  
Db 72 TGGGCGCTACGGCAGCGCCACTGCTCCCAAGACGACATGACCGCTCTCTCGGGGAC 131  
QY 41 TyrGlyGlnThrSerIleMetSerProArgValProGlnInaSPTrIleuSerGlnPro 60  
Db 132 TATGACAGACATCCAAAGATGAGCCGCGCGTCCCTCAGCAGAGACTGTGCTCAACCC 191  
QY 61 ProAlaArgValThrIleYsgIAlaIeuserValSerGlnuSpGInserleuPheGlu 80  
Db 192 CCAGCCAGGCTCACCTCAAGATGATGATACCTCAACAGGCTCAAGGAC 251  
QY 81 SerProAspPcysSerValAlaIaTyrGlyGlyMetValSerSerSerAspAsnVal 100  
Db 252 TCCCTGATGATGTCACGATGCGCAAGGTGGAGATGTGGGACGCCCGGACACCGCTC 311

QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThr 120  
Db 312 GGGATGAACATACAGACACTACATGAGAGAGACATGCGACACCCCAACATGACACACC 371  
QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuThrSerThrAspHisValArg 140  
Db 372 AACGAGCGCGCGTCATCGTCCCTGAGATCCACGCTATGAGATACAGACATGTCCGG 431  
QY 141 GlnTrpLeuGluTrpAlaValIleYsgIAlaIeuserValSerGlnuSpGInserleuPhe 160  
Db 432 CAGTGGCTGGAGTGGCGAGTCACAAAGATACGGGCTTCGGACGTGACATCTATTATTC 491  
QY 161 GlnAsnIleAspGlyIleuGluCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
Db 492 CAGATATTCACGCGGAGGAGAGCTGTGCATGACCAAGATGACTTCCAGAGGCTCACC 551  
QY 181 ProSerTyrAsnAlaAspIleuLeuSerHisIleuHis 193  
Db 552 CCGAGCTACAAAGCTGACATCTTGTGTGCGACCTCCAC 590  
RESULT 10  
BG390291 915 bp mRNA linear EST 12-MAR-2001  
LOCUS BG390291  
DEFINITION 602416255F1 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:4524852 5',  
mRNA sequence.  
ACCESSION BG390291  
VERSION BG390291.1 GI:13283739  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 915)  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
UNPUBLISHED (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: c9abps-remail.nih.gov  
Tissue Procurement: ARCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LHAM10429 row: m column: 13  
High quality sequence stop: 820.  
FEATURES  
source Location/Qualifiers  
1..915  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:4524852"  
/clone\_1lb="NIH\_MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: PCMV-SPORT6; Site.1: NotI;  
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC library."  
BASE COUNT 246 a 252 c 227 g 190 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.53e-89 Length: 915  
Score: 975.50 Matches: 215  
Percent Similarity: 77.36% Conservative: 14  
Best Local Similarity: 72.64% Mismatches: 42  
Query Match: 39.91% Indels: 28  
DB: 12 Gaps: 5  
US-09-902-772-2 (1-451) x BG390291 (1-915)

```

QY      1 MetAlaSerThrIleuylSGuAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      59 ATGGCCAGCAGCTATTAAAGAGCCCTTAATGCTGTGTGATGAGAGCAGCTGTTGTTGAG 118
QY      21 CysAlaTyrGlySerProHISLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40
DB      119 TGTGGCTACGGAAGCCACACCTGGCTAAGACAGAGTGCACCGCTCTCTCTCCAGCGAC 178
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPheLeuSerGlnPro 60
DB      179 TATGACAGACTTCCAGATGAGAGCCAGCGCTCCCTCAGCAGATGGCTGTCTCAACCC 238
QY      61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB      239 CCAAGCAGGGTCCACCATCAAAATGAATGATACCCCTAGCCAGGTGAATGGCTCAAGGAC 298
QY      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB      299 TCTCCTGATGAATGCACTGTGGCCAAAGGGGAGAGTGGTGGGACCCAGACCCGTT 358
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleLeuProProAsnMetThrThr 120
DB      359 GGGATGAATACGGCAGCTACATGAGAGAGACACATGCCACC-CCAACATGACACAGC 417
QY      121 AsnGluArgArgValIleValProAla-AspProThrLeuTyrSerThrAspHisValArg 140
DB      418 AACGAGCGCAGAGTATACGTCGCAAGATCCTACCTATGAGTACACACATGTCGCG 477
QY      140 GgIntPrLeuGluTyrAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB      478 GCAGTGGCTGGAGTGGGGCGGTGAAGAATATGGCTTCCAGAGCTCAACATCTTGTATT 537
QY      160 eGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB      538 CCAGAACATCGATGGAGAGAACTGTGCAGATGACCAAGAGACGACTTCCAGAGGCTCAC 597
QY      180 rProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgLysArg---G1 199
DB      598 CCCAGAGTACAAATGCGACAT-CTTCTCTCAATCTCTCACTACCTCGAAGACTCCCTTC 656
QY      199 yAlaThrPheIlePhePro-----AsnThrSerValTyrProGlnAlaThrGlnArg11 217
DB      657 TTCACATTTGACTTCAGATGATGTTGATTAAAGCTTCAAAACTCCACGGTTATGTC 716
QY      217 eThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrThrSerHisSe 237
DB      717 ATGCTAGAAACACAGATTTACATATGAGCCCCCAGAGAGATCAAGCTGGACGTCACGG 776
QY      237 rHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGlu-- 256
DB      777 -CACCCACGCCAGTCACAGCTGTACATCTCTTCCAGCTGCCAAATGAGACAGCTCT 835
QY      257 -----AspGlnArgProGlnLeuAspProTyr 265
DB      836 CGGAAATCTATGATTTCTGGCACAGTTCGCTGCATCGGGGGCGACCTTGGATGCC-- 893
QY      265 rGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
DB      894 -----GGGCGGGGAGC-----CCCGGCTCC 914

RESULT 11
LOCUS   AUI36709 701 bp mRNA linear EST 02-AUG-2002
DEFINITION AUI36709 PLACE1 Homo sapiens cDNA clone PLACE1004911 5', mRNA
sequence.
ACCESSION AUI36709
VERSION   AUI36709.1 GI:10997248
KEYWORDS  EST.
SOURCE    human.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 701)
AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
           Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
           Isogai,T.
TITLE      HRI human cDNA project
JOURNAL    Unpublished (2000)
COMMENT    Contact: Takao Isogai
           Genomics Laboratory
           Helix Research Institute
           1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           Tel: 81-438-52-3975
           Fax: 81-438-52-3986
           Email: genomics@hri.co.jp
           HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
           Research Institute; cDNA library construction: Department of
           Virology, Institute of Medical Science, University of Tokyo, and
           Helix Research Institute.
FEATURES   Location/Qualifiers
           source          1..701
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="PLACE1004911"
                        /clone_1lb="PLACE1"
                        /tissue="placenta"
                        /note="Vector: pME18SF13"
BASE COUNT 199 a 185 c 166 g 145 t 6 others
ORIGIN
Alignment Scores:
Pred. No.: 1,62e-88 Length: 701
Score: 971.50 Matches: 185
Percent Similarity: 94.15% Conservative: 8
Best Local Similarity: 90.24% Mismatches: 10
Query Match: 39.75% Indels: 2
DB: 9 Gaps: 1

US-09-902-772-2 (1-451) x AUI36709 (1-701)
QY      1 MetAlaSerThrIleuylSGuAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      85 ATGGCCAGCAGCTATTAAAGAGCCCTTAATGCTGTGTGATGAGAGCAGCTGTTGTTGAG 144
QY      21 CysAlaTyrGlySerProHISLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40
DB      145 TGTGGCTACGGAAGCCACACCTGGCTAAGACAGAGTGCACCGCTCTCTCTCCAGCGAC 204
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPheLeuSerGlnPro 60
DB      205 TATGACAGACTTCCAGATGAGAGCCAGCGCTCCCTCAGCAGATGGCTGTCTCAACCC 264
QY      61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB      265 CCAAGCAGGGTCCACCATCAAAATGAATGATACCCCTAGCCAGGTGAATGGCTCAAGGAC 324
QY      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB      325 TCTCCTGATGAATGCACTGTGGCCAAAGGGGAGAGTGGTGGGACCCAGACACCGCTT 384
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleLeuProProAsnMetThrThr 120
DB      385 GGGATGAATACGGCAGCTACATGAGAGAGACATGTCACCCCAACATGACACAGC 444
QY      121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140
DB      445 AACGAGGCGCAGAGTATACGTCGCAAGATCTTACGCTATGAGTACAGACATGTGCGG 504
QY      141 GgIntPrLeuGluTyrAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB      505 CAGTGGCTGGAGTGGGGGGGTGAAGAATATGAGCTTCCAGACGTCACATCTTGTATTTC 564
QY      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThr-TyrAspAspPheGlnArgLeuThr 180
DB      565 CAGAACATCATGTGGAGAAAGAACTGTGCAAGATGACCAAGAGACGACTTCNAGANGCTCAA 624

```

OY 180 PProSerTyraAlaAspIleLeuSerHisLeuHisTyrLeuArgGluArgGlyAl 200  
 Db 625 CCCAGCTACACCCGACATCTCTCTCCTCAGATCTCAGTACTC---AAGAGAGACTC 681  
 OY 200 aThrPheIlePhe 204  
 Db 682 CTCCTTCACATTT 694  
 RESULT 12  
 AM948986 629 bp mRNA linear EST 31-MAY-2000  
 LOCUS QV4-FR0005-110500-201-F02 FR0005 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION AM948986  
 ACCESSION AM948986.1 GI:8126760  
 VERSION  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 629)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 ,M.J., Soares,F., Brentani,R.K., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 COMMENT 20202663  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV4-FR0005-110  
 500-201-f02&t3=2000-05-11&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 628.  
 Location/Qualifiers  
 1..629  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="FR0005"  
 /dev\_stage="Adult"  
 /note="Organ: prostate,tumor; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."  
 BASE COUNT 166 a 173 c 171 g 119 t  
 ORIGIN  
 Alignment Scores:  
 Pred. NO.: 1.96e-88 Length: 629  
 Score: 970.00 Matches: 180  
 Percent Similarity: 97.40% Conservative: 7  
 Best Local Similarity: 93.75% Mismatches: 5  
 Query Match: 39.69% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-902-772-2 (1-451) x AM948986 (1-629)  
 OY 6 LysGluLeuSerValIleSerGluAspGlnSerLeuPheGluGlyAlaTyrGlySer 25  
 ::

Db 19 CAGGAGCCCTTATGACTGTGAGTGAGGACCAGTCGTTGTTGAGTGGCTACGGAAC 78  
 OY 26 ProHisLeuAlaTyrThrGluMetThrAlaSerSerSerSerGlyGlnThrSer 45  
 Db 79 CCACACCGGCTAAGACAGAGATGACCGCGTCTCTCCACGACTATGACAGACTTCC 138  
 OY 46 LysMetSerProAlaGValProGlnGlnAspTyrLeuSerGlnProProAlaArgValThr 65  
 Db 139 AAGATGAGCCACCGCCGCTCCACAGAGATTGGCTGTCTCAACCCACGACGAGGTACC 198  
 OY 66 IleLysMetGluGlyAsnProAsnGlnValAsnGlySerAlaGlnSerProAspAspCys 85  
 Db 199 ATCAAAATGAAATATTAACCTACGACGTAATGCTCAAGAACTCTCTGATGAATGC 258  
 OY 86 SerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGly 105  
 Db 259 AGTGTGGCCAAAGCGGGAAGATGGTGGCGACGCCAGACACCGTTGGATGAATACGGC 318  
 OY 106 SerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsnGluArgVal 125  
 Db 319 AGCTACATGGAGGAGAGACATGTCACCCCAACATGACACGACGACGACGAGATT 378  
 OY 126 IleValProAlaAspProThrLeuTyrSerThrAspHisValArgGlnTyrLeuGluTyr 145  
 Db 379 ATCGTCCAGCAGATCTCTACGCTATGAGTACAGACCATGTGCGGACGTGGTGGATGG 438  
 OY 146 AlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGly 165  
 Db 439 GCGGTGAAGAATATGCGCTTCCAGACGTCATCTGTATTCAGAAATCATGATGGG 498  
 OY 166 LysGluLeuGlyLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAla 185  
 Db 499 AAGGAACTGTCAAGATGACCAATGACGACTTCCAGAGGCTCCACCCACGCTACACGCC 558  
 OY 186 AspIleLeuSerHisLeuHisTyrLeuArgGlu 197  
 Db 559 GACATCCTCTCTCATCATCTCCACTACTAGAGAG 594  
 RESULT 13  
 B0953977 1074 bp mRNA linear EST 21-AUG-2002  
 LOCUS AGNCCOURT 8877989 Lupski\_sciatic\_nerve Homo sapiens CDNA clone  
 DEFINITION IMAGE:6198592 5', mRNA sequence.  
 ACCESSION B0953977  
 VERSION B0953977.1 GI:22369455  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1074)  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: L1AM13610 row: h column: 17  
 High quality sequence stop: 575.  
 Location/Qualifiers  
 1..1074  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="6198592"  
 /clone\_lib="Lupski\_sciatic\_nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"



K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y., and Hayashizaki, Y.  
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Funct. Genomics 2 pre, L72-L86 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

# FEATURES

Location/Qualifiers

1..675  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="A630068021"  
/clone\_lib="RIKEN full-length enriched, 3 days neonate thymus"  
/tissue\_type="thymus"  
/dev\_stage="3 days neonate"  
/lab\_host="DH10B"

/note="Site\_1: Salt; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTCVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGAGTTATTAATTATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda phage I."

BASE COUNT 186 a 165 c 176 g 148 t  
ORIGIN

## Alignment Scores:

Pred. No.: 3.8e-86 Length: 675  
Score: 948.00 Matches: 177  
Percent Similarity: 96.84% Conservative: 7  
Best Local Similarity: 93.16% Mismatches: 6  
Query Match: 38.79% Indels: 0  
DB: 10 Gaps: 0

US-09-902-772-2 (1-451) x BB639043 (1-675)

QY 1 MetAlaSerThrIleuValGluAlaLeuSerValValSerGluInsPheGlu 20  
DB 106 ATGCCCGACATATTAAAGAGGCTGTCTGAGTGTGAGCGAGACAGTCACTATTGAG 165  
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
DB 166 TGTGCTACGGAAGCCACACCTGTGTAAGACAGATGACCCGATCTCTCCAGGAC 225  
QY 41 TyrGlyIleuThrSerLysMetSerProAlaValProGlnGlnAspTrpLeuSerGlnPro 60  
DB 226 TATGCGCAGACATCCAGATGAGTCCAGAGTCCCTCAGCAGGAGCTGGCTGTCAAGCC 285  
QY 61 ProAlaArgValThrIleLysMetGluCysAspProAsnGlnValAsnGlySerArgAsn 80  
DB 286 CCAAGCCAGGCTCACCACCAAGATGAGTGCACCCCTGAGTGAATGTTCCAGGAC 345  
QY 81 SerProAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
DB 346 TCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 405  
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120  
DB 406 GGGATGAGTACGCGACCTACATGAGAGAGAGATGATGATGATGATGATGATGATGAT 465  
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140

DB 466 AATGAGCGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525  
QY 141 GlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAlaIleLeuPhe 160  
DB 526 CAGTGGCTGAGTGGCGGCGTGAAGAATATGCGCTCTCCATGATGACGCTTACTATT 585  
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
DB 586 CAGATATTCGATGCGAGAGAGCTGTGCAAGATGACAAAGATGATGATGATGATGATGATGAT 645  
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSer 190  
DB 646 CCGAGCTACATGCGGACATTTCTCTCA 675

RESULT 15  
BM489636  
LOCUS  
DEFINITION  
634 bp mRNA linear EST 07-FEB-2002  
pgm2n.pk011.124 Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA  
clone pgm2n.pk011.124 5' similar to sp|Q90837|ERG.CHICK  
TRANSCRIPTIONAL REGULATOR ERG pIR1560754 transcription factor erg - chicken emb|CAA54404.1| (X77159) ERG [Gallus gallus], mRNA sequence.  
BM489636  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
chicken.  
Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Arctosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
1 (bases 1 to 634)  
Codburn, L.A. and Monson-O'Oran, E.  
ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IRAFs Animal Genome Project  
Unpublished (2002)  
Contact: Larry A. Codburn  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: codburn@udel.edu, www.chickest.udel.edu.

## JOURNAL COMMENT

Unpublished (2002)  
Contact: Larry A. Codburn  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: codburn@udel.edu, www.chickest.udel.edu.

## FEATURES

Location/Qualifiers

1..634  
/organism="Gallus gallus"  
/strain="Commercial broiler and Ottawa Res. Centre strains 90 & 21"  
/db\_xref="taxon:9031"  
/clone="pgm2n.pk011.124"  
/clone\_lib="Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n)"  
/sex="Male and Female"  
/tissue\_type="Breast muscle, leg muscle and epiphyseal growth plate"  
/dev\_stage="Breast, leg: Embryo(d19); post-hatch(1d, 1.3, 5, 7, 9, 11 weeks); growth plate(1d, 7d, 14d post-hatch)"  
/lab\_host="E. coli EMD10B"  
/note="Vector: PCMVSPORT6; library made from equivalent pools of total RNA isolated from each tissue (embryonic muscle 33.3%, juvenile muscle 33.3%, and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"

## BASE COUNT

192 a 164 c 138 g 140 t

## ORIGIN

Alignment Scores:  
Pred. No.: 3.88e-86 Length: 634  
Score: 947.50 Matches: 182  
Percent Similarity: 86.67% Conservative: 0  
Best Local Similarity: 86.67% Mismatches: 1  
Query Match: 38.77% Indels: 27  
DB: 13 Gaps: 1

US-09-902-772-2 (1-451) x BM489636 (1-634)

```
QY      120 ThrAsnGluArgValIleValProAlaSPProThrLeuTrpSerThrAspHisVal 139
      |||
Db      3 ACCAATGAAAGAGATTATGTGTCAGAGATCTTCAGTTATGTAGACACAGACCATGTA 62
      |||
QY      140 ArgGlnTrpLeuGlnTrpAlaValIleGluTyrGlyLeuProAspValAspIleLeu 159
      |||
Db      63 CGGAGTGGCTGGAGTGGAGTGAAGAGATGTCTTCAGACGTGACATCTGTG 122
      |||
QY      160 PheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeu 179
      |||
Db      123 TTCAGAAACATGTGATGGAAAGAGTGTGTAATAATGACCAAGATGACTTCAGAGACTC 182
      |||
QY      180 ThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg--- 198
      |||
Db      183 ACGCGAGCTATAACGAGATATCTCTGTCAACACTACACTACACTACAGAGACTCT 242
      |||
QY      198 ----- 198
      |||
Db      243 CTTCACATTGACTTCAGATGATGTGATTAAGGCTTACAAAACCTCCACGGTTATG 302
      |||
QY      199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGlu 212
      |||
Db      303 CATGCTGAACACACAGAGAGCCACTTTATTTTCCAAATACATCAGTTTACCCAGAA 362
      |||
QY      213 AlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAla 232
      |||
Db      363 GCAACGCCAAGAATTAACACAGGCCAGATTACCTTATGAGCAAGCGAGAGATCAGCG 422
      |||
QY      233 TrpThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrVal 252
      |||
Db      423 TGGACGAGTCACAGCCATCCACTCACTCAAAAGCTACCAACCATCATCTCAACAGTG 482
      |||
QY      253 ProLysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
      |||
Db      483 CCCAAAACAGAACACAGCCGCTCTCAGATTGATCTTATCAGATTCTTGACGACGACG 542
      |||
QY      273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeu 292
      |||
Db      543 AGCCGCTCTGCAAAATCCAGGAGTGGGAGATACAGCTATGGCAGTCCCTACTGAGCTT 602
      |||
QY      293 LeuSerAspSerSerAsnSerAsnCysIle 302
      |||
Db      603 CTGTGGAGAGCTCCAACTCCAACTGCATC 632
      |||
```

Search completed: July 28, 2003, 07:39:12  
Job time : 1906.27 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:14:59 ; Search time 3810.11 Seconds

(without alignments)  
3651.115 Million cell updates/sec

Title: US-09-902-772-4  
Perfect score: 2588  
Sequence: 1 MASTIKKALSVSEDSQSLFE.....IYPNTRLPAAHMPSHLCTYY 478

Scoring table: BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPFO.spool/US09902772/runtat\_23072003\_093657\_14878/app\_query.fasta\_1.1294  
-DB=GenEmbl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09902772.GCgn\_1\_1\_4707\_etruntat\_23072003\_093657\_14878 -NCPV=6 -ICPV=3  
-NO\_MAP -LARGESUBSTRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_da:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_inu:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pla:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2436	94.1	1516	5 GGERG	X77159 G.gallus ER
2	2436	94.1	1528	6 E31254	E31254 Protein hav
3	2293.5	88.6	2209	10 AB073080	AB073080 Mus muscu
4	2262.5	87.4	1413	5 AY065661	AY065661 Gallus ga
5	2262.5	87.4	1447	6 E31253	E31253 Protein hav
6	2186	84.5	2133	10 AB073079	AB073079 Mus muscu
7	2180.5	84.3	3166	9 HUMERG2	HUMERG2
8	2180	84.2	1798	5 XLAJ4125	XLAJ4125 Xenopus l
9	2177.5	84.1	1808	10 AB031088	AB031088 Rattus no
10	2149.5	83.1	2131	10 AB073078	AB073078 Mus muscu
11	2015.5	77.9	2012	5 XLAJ4126	XLAJ4126 Xenopus l
12	1808.5	69.9	3126	9 HUMERG11	M21535 Human erg p
13	1491.5	57.6	1436	5 DRE249590	AJ249590 Danio rer
14	1482	57.3	1359	9 AY029368	AY029368 Homo sapi
15	1482	57.3	2394	9 BC010115	BC010115 Homo sapi
16	1482	57.3	2916	6 A36461	A36461 Sequence 2
17	1482	57.3	2938	6 AR080101	AR080101 Sequence
18	1482	57.3	2938	9 HSHDMFLI	X67001 H.sapiens H
19	1482	57.3	2957	9 HUMERGBFLI	M98833 Homo sapien
20	1482	57.3	2959	9 BC001670	BC001670 Homo sapi
21	1482	57.3	6849	12 AY029367	AY029367 Synthetic
22	1479	57.1	1673	9 S45205	S45205 Flt-1-Frien
23	1474	57.0	1729	10 MPEFLI	X59421 Mouse Flt-1
24	1440.5	55.7	4403	5 XLEFLG	X66979 X.laavis MR
25	1432	55.3	3490	5 CCRNAFLI	Y14773 Coturnix co
26	1420	54.9	3545	5 CCEFLONCO	Y14774 Coturnix co
27	1394	53.9	1932	9 HUMFLITIA	M93255 Human FLI-1
28	1341	51.8	1245	5 AE177538	AE177538 Danio rer
29	1253.5	48.4	816	10 S66169	S66169 Erg-3=Immun
30	1036.5	40.1	762	9 S72621	S72621 EMS...erg f
31	834	32.2	1431	9 AFE327066	AFE327066 Homo sapi
32	809.5	31.3	14518	9 AP001423	AP001423 Homo sapi
33	809.5	31.3	100000	9 AP000021	AP000021 Homo sapi
34	809.5	31.3	100000	9 AP000163	AP000163 Homo sapi
35	809.5	31.3	340000	9 AP001731	AP001731 Homo sapi
36	778	30.1	588	9 S72865	S72865 EMS...EMS-e
37	702.5	27.1	66352	9 HSY17293	HSY17293 Homo sapien
38	702.5	27.1	187131	2 AP001535	AP001535 Homo sapi
39	702.5	27.1	217382	2 AP001122	AP001122 Homo sapi
40	604	23.3	2180	3 AY060316	AY060316 Drosophila
41	602	23.3	634	3 SULERG	M61067 Sea urchin
42	514.5	19.9	491	3 HDI311813	AJ311813 Hediaste d
43	485.5	18.8	169741	9 AC097468	AC097468 Homo sapi
44	484.5	18.7	83830	2 AC095777	AC095777 Rattus no
45	484.5	18.7	192968	2 AC127107	AC127107 Rattus no

RESULT 1

## ALIGNMENTS

GGERG LOCUS 1516 bp mRNA linear VRT 27-APR-1995  
 DEFINITION G.gallus ERG mRNA.  
 ACCESSION X77159  
 VERSION X77159.1 GI:790439  
 KEYWORDS ERG gene.  
 SOURCE chicken.  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 1516)  
 Dhoradin, P., Dewitte, F., Desbiens, X., Stehelin, D. and  
 Dutergue-Coquillaud, M.  
 Mesodermal expression of the chicken erg gene associated with  
 precartilaginous condensation and cartilage differentiation  
 Mech. Dev. 50 (1), 17-28 (1995)  
 JOURNAL MEDLINE 95329425  
 PUBMED 7605748  
 2 (bases 1 to 1516)  
 Dutergue-Coquillaud, M.  
 Direct Submission  
 Submitted (17-JAN-1994) M. Dutergue-Coquillaud, CNRS UA 1160,  
 Oncologie Moleculaire, Institut Pasteur, 1 rue Calmette, 59019  
 Lille, FRANCE  
 FEATURES  
 source Location/Qualifiers  
 1. 1516  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"  
 /clone\_lib="lambda gt10"  
 /dev\_stage="adult"  
 63. 1499  
 /gene="ERG"  
 63. 1499  
 /gene="ERG"  
 /codon\_start=1  
 /protein\_id="CA54404.1"  
 /db\_xref="GI:790440"  
 /translation="MASTIKKALSVSESDSLFCAYASPHLAKTENTASSSEYQOT  
 SKNSPRVODDMISOPPARVYIKMECPNNOVNGSRNSPDOSVAKGKSSNVNVM  
 NYGSMERKHIPPNMTNERRYIVPADPTLMSDHRWLEFAVKEKYGJPDNDILF  
 ONIDKRLCKMTKDDPRLTPSYNAOILLSHLRLRTPPLHLSDVDYALQNSPRL  
 MHAANTGATFIFPNISVPEATORTTRDLPYEQARSAWTSHPPTOSKATOPSS  
 STYKTEDDQRPOLDPYQILGPTSSRLNPSGOIQLOFLELSDSNSNCITWECT  
 NGERKMTDPDEVARWGERKSKPMNADKLSRALRYYYDKNIMTKVHGKRAYKFDH  
 GIAOALOPHPESSEMYKYPSPDLPYMSSTYHAHPQKMFVAVHPALPYTSSSFAAPNP  
 YMSPTGCIPTNRLPAHMPHSHLGTYY"  
 BASE COUNT 458 a 392 c 327 g 339 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,45e-158 Length: 1516  
 Score: 2436.00 Matches: 458  
 Percent Similarity: 91.97% Conservative: 0  
 Best Local Similarity: 91.97% Mismatches: 0  
 Query Match: 94.13% Indels: 40  
 Gaps: 2  
 US-09-902-772-4 (1-478) x GGERG (1-1516)  
 QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 63 ATGCAGACACTATTAGAGAGCATATACAGTGTGATGAGTGAAGACAGCTCTGTTGAG 122  
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40  
 Db 123 TGTGCTACGAGATGCCCACTTGCAAGACAAAGATGACAGAGCTCTTCACAGTAA 182  
 QY 41 TyrGlyInThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGluPro 60  
 Db 183 TATGGCAAAACATCAAAAGATGAGCGCGGCTTCCCAAGACAGAGACTGTTATCACAGCCC 242

QY 61 ProAlaArgValIrrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 243 CCGGCCAGAGTTTACCATTAAGATGAGAGTAAACCAACCAAGTAAATGAGGTCAAGAAAT 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnVal 100  
 Db 303 TCACCTGATGACTCAGCGGTGGCAAAAGAGAGAAATGTTTACAGCTTCAGACAAATGTT 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProPropoAsnMetThrThr 120  
 Db 363 GGGATGAACCTATGGAAGCTACATGGAAGAGAGATATTCGGCTCCAAATATACAAAC 422  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140  
 Db 423 AATAAGAGAGAGTAAATGTTGCCAGAGATCTCAGTTATGAGACAGACAGATGACGG 482  
 QY 141 GluTrpLeuGluTrpAlaValAlaLysGluTyrGlyLeuProAspValAspIleLeuPhe 160  
 Db 483 CAGTGGCTGGAGTGGAGTGAAGAGATAGTGTCTCCAGACGTGACATCTGTGTTC 542  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 543 CAGAACATGATGGGAAAGAGTGTGTAAATGACCAAGATGACTTCCAGAGACTGACG 602  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluTrpProLeu 200  
 Db 603 CCGAGCTATTAACGAGATATCTCTGTACACCTACATACCTCAGAGAGACTCTCTT 662  
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuLnsnsrProArgLeuMetHis 220  
 Db 663 CCACATTTGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722  
 QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluVala 240  
 Db 723 GCTAGAAACACAGAGAGAGGACACTTATTTTCCAAATACATGAGTTTACCAGAGCA 782  
 QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 260  
 Db 783 ACGCAAGATTAACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 842  
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280  
 Db 843 ACGAGTACAGCCATCCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 902  
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300  
 Db 903 AAAACAGACAGCAGCGTCTCAGTTAGATCCTTATCAGATTCGTGACGACGACGACG 962  
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeuLeu 320  
 Db 963 CGTCTTCAAAATCCAGGAGTGGGAGATACAGTATGAGCAGTTCCTACCTGAGCTTCTG 1022  
 QY 321 SerAspSerSerAsnSerAsnGlyIleThrTrpGluLysThrAsnGlyLysPheLysMet 340  
 Db 1023 TCGGACAGCTCCAACTCACTCACTCACTGAGGAGGCAAAATGGGAGTTCAAGATG 1082  
 QY 341 ThrAspProAspGluValAlaArgArgTrpGlyLysLysSerLysProAsnMetAsn 360  
 Db 1083 ACAGACCCCTATGATAGTGGCTCGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1142  
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380  
 Db 1143 TATGACAAACTCAGCCGTGACCTGCTACTACTATGACAAAATATATAGATAAAGTT 1202  
 QY 380 ----- 380  
 Db 1203 CATGTAAACGCTATGCTTCAAAATTTGATTTCCACGGAATTCCTCAGCCCTCCAGCTT 1262  
 QY 381 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400  
 Db 1263 CACCTTCACAAATCATCATCATGATCAATATACCATTCAGAGACCTCCCTACATGATTCCTAC 1322

QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420

Db 1322 ----- 1322

QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440

Db 1323 CATGACACACCCAGCAAGATGAATCTTGTAGCTCCCATCCCGCTTGTGCGGTAACC 1382

QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThgIlyIleTyr 460

Db 1383 TCATCCAGCTTTTGTGCTGCCCCCTATCCATCTGGAATTCACCACTGGAGCATCTAC 1442

QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleThrTyrTyr 478

Db 1443 CCCAATACCAAGCTGCACTGCTCATATGCTTCCATCTTGCGACCTACTAC 1496

RESULT 2

E31254 1528 bp DNA linear PAT 18-JUN-2001

LOCUS E31254

DEFINITION Protein having cell calcifying inhibitory activity and gene encoding the same.

ACCESSION E31254

VERSION E31254.1 GI:13025686

KEYWORDS JP 199075871-A/2.

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1528)

AUTHORS Hiroyasu, I., Yoshinobu, H., Martijio, P., Joel, R. and Helena, E.

TITLE Protein having cell calcifying inhibitory activity and gene encoding the same

JOURNAL Patent: JP 199075871-A 2 23-MAR-1999;

CHUGAI PHARMACEUT CO LTD, HE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA

COMMENT OS Unidentified

PN JP 199075871-A/2

PD 23-MAR-1999

PE 29-MAY-1998 JP 1998166076

PR 18-JUN-1997 US 08/878177 20-JUN-1997 US 60/050297 PI HIROYASU IMAMOTO, YOSHINOBU HIGUCHI, MARIJITO PASHIPIKI, PT JOEL ROZENBROOD, E

PI HELENA E

PC C12N15/09,A61K48/00,C07K14/465,C07K16/18,C12Q1/68//A61K38/00, PC C12P21/02,

PC C12P21/08,C12N15/00,A61K37/02

CC Strandedness: Single;

CC Topology: Linear;

FT source 1.1528

FT 1.1528

FEATURES

source location/Qualifiers

1.1528 /organism="unidentified"

/db\_xref="taxon:32644"

BASE COUNT 464 a 395 c 330 g 339 t

ORIGIN

Alignment Scores:

Pred. No.: 2,48e-158 Length: 1528

Score: 2436.00 Matches: 458

Percent Similarity: 91.97% Conservative: 0

Best Local Similarity: 91.97% Mismatches: 0

Query Match: 94.13% Indels: 40

Db: 6 Gaps: 2

US-09-902-772-4 (1-478) x E31254 (1-1528)

QY 1 MetaLaserThrlLysGlnAlaLeuSerValIserGluAspGlnSerLeuPheGlu 20

Db 63 ATGGCAGACACTATTAAGGAACATTATCAGTGGTGAAGGAGACAGCTCTGTTTGAG 122

QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40

|||||

Db 123 TGTGCTACGAGATGCCCCACCTTGCAAGACAGAATGACAGCCTCTTCCAGTGA 182

QY 41 TyrGlyGlnThrSerLysMetSerProAlaValProGlnGlnAspTyrLeuSerGlnPro 60

Db 183 TATGGGCAACATCAAGATGAGCCCGCGCTTCCACAGACAGTGGTATCACAGCCC 242

QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80

Db 243 CCGGCCAGAGTTACCATTTAGATGGAGTGAACCCCAACCAGGTTAATGGCTCAAGGAAT 302

QY 81 SerProAspAspCysSerValAlaLysGlyIlyLysMetValSerSerSerAspAsnVal 100

Db 303 TCACCTGATGATGCAGCGGTGGCAAAAGAGGGAATAATGTTAGCATTCACACATGTT 362

QY 101 GlyMetAsnTyrGlySerTyrMetGluGlyLysHisIleProProAsnMetThrThr 120

Db 363 GGGATGAACTATGAGAGCTACATGAGAGAGAGCATATTCCTCCCAATATGACAMCC 422

QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140

Db 423 AATGAAACGAAGATTTATTTGCTGACAGATCTTACGTCATGACACACATGTACCG 482

QY 141 GlnTyrPLeuGluTyrPalAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160

Db 483 CAGTGGCTGAGTGGGAGCTGAGAGGATGTGCTTCCAGAGCTGGACATCTTGTGTTTC 542

QY 161 GlnAsnIleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheIleArgLeuThr 180

Db 543 CAGAACATGATGAGGAAAGATGTGTAAAGACCAAGAGATCTCCAGAGACTCAGC 602

QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200

Db 603 CCGAGCTAATACGACGATATCTCTCTGTACACCTACACTACTACGAGAGACTCTCTT 662

QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220

Db 663 CCACATTTGACTTTCAGATGATGTGATAAGGCTTACAAACCTCCACAGGTTATGATCAT 722

QY 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240

Db 723 GCTAAGAACACGAGAGAGCCACTTTATTTTCCAAATATCATGATTTATCCACAGAAGA 782

QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlyGlnAlaArgArgSerAlaTyr 260

Db 783 ACGCAAGACATATACACAAAGCCAGATTTACTTTAGACAAAGCAGAGATCAGCGTGG 842

QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280

Db 843 ACGAGTCAACAGCATCCCTCAGTCAAAAGCTACCAACCATCATCTTCAACAGCTGCC 902

QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300

Db 903 AAAACAGAGACCAACGCTCTCTCAGTTAGATCTTATCAATCTTGGAGCCAGCAGCAC 962

QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPGLnPheLeuLeuLeuLeu 320

Db 963 CGCTTTCGCAATCCAGGAGATGGGACAGATACAGCATGTGCACTTCTACTGAGCTTCTG 1022

QY 321 SerAspSerSerAsnSerAsnCysIleThrTyrGluGlyThrAsnGlyLysLeuLeuMet 340

Db 1023 TCGGACACTCTCAATCCAACTGCAATCCTGAGAGGCGCAAAATGGGAGATTCAAAGAT 1082

QY 341 ThrAspProAspGluValAlaArgArgTyrPGLnGlyLysSerLysProAsnMetAsn 360

Db 1083 ACAGACCTGATGAGAGTGGCTGGCTGGGAGAGAGAGAAAGCAAACTTAACATGAGAC 1142

QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380

Db 1143 TATGCAAACTAGCGCGTGGCACTTGTGCTACTACTATGCAAAATATTTATGCTAAAGTT 1202

QY 380 ----- 380

Db 1203 CATGCTAAACGCTATGCTACAAATTTGATTTCCACGGAATGCTCAGGCTCCAGCCT 1262

OY		381	HISPRPGLGUSerSerMetTylusTyProserAspLeaProTyrMeSerSerTyr	400
DB		1263	CACCTCCAGATCATTCATGTACAAATACCACATCAGCCCTCCCTACTATAGTTCTCAC	1322
OY		401	HISgLySarGTYrAlatYrLysPheaspPheHISgLYlLaGlalalaugLnPro	420
DB		1322	-----	1322
OY		421	HISlaHISproGINlysmetAsnPheValAlaProHISproPROlalaleuProvalThr	440
DB		1323	CATGCACACCCCCAGAAAGATGAACCTTGTTACTCTCCCAATCCCCTGGCTTCCCGTAACC	1382
OY		441	SerSerSerPhePheAlaAlaProasnProTyrTrpaSnSerProThIGLYIyleTyr	460
DB		1383	TCAITCCAGCTTTTTTGGCGCCCTAATGCCAATCTGGAATTACCAACAATGGAGCATCTAC	1442
OY		461	ProASnthArgLeuProAlaAlaHisMetProSerHisLeuGIyThrTyrTyr	478
DB		1443	CCCAATACACGAGCTGCAGCTGTCATATATGCTTCCCATCTTGGCACTACTAC	1496
RESULT 3				
AB073080				
LOCUS		2209 bp	mRNA	linear
DEFINITION			Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG,	
ACCESSION		AB073080		
VERSION		AB073080.1	GI:16197544	
KEYWORDS				
SOURCE				
ORGANISM			Mus musculus cDNA to mRNA, clone:Erg-3.	
REFERENCE			Mus musculus	
AUTHORS			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y. Mus musculus Erg mRNA Published only in Database (2001) 2 (bases 1 to 2209) Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y. Direct Submission Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Shuto-chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://ngp.gsc.riken.go.jp/, Tel.:81-45-503-9111, Fax:81-45-503-9170)	
TITLE	JOURNAL			
REFERENCE	AUTHORS			
FEATURES	SOURCE			
	gene			
	CDS			
BASE COUNT		606 a	608 c	545 g 450 t
ORIGIN				

Alignment Scores:	Pred. No.:	2,34e-148	Length:	2209
Score:	2293.50		Matches:	431
Percent Similarity:	88.76%		Conservative:	11
Best Local Similarity:	86.55%		Mismatches:	15
Query Match:	88.62%		Indels:	41
DB:	10		Gaps:	3
US-09-902-772-4 (1-478) x AB073080 (1-2209)				
QY	2	AlAserrthrllelysgluAlaleuSerValValSerGluAspGlnSerLeupheglucys	21	
DB	172	GCAGCTCATATTAAAGAGAGCCCTTGTCAGTTGTGACGAGAGACAGCAGCTATTGTAGTGT	231	
QY	22	AlAtyrGlySerProHisleuAlalysrThrGluMetThrAlaSerSerSerGlyuIyr	41	
DB	232	GCCTAAGGAGAGCCACCTGGCTAAGACAGAGATGACCGCATCTTCCTGAGTACTAT	291	
QY	42	GlyGlnThrSerIyMetSerProArqValProGlnGlnAspThrIleuSerGlnProPro	61	
DB	292	GGCCGACATCCAAAGATGATGCCAGAGTCCCTCAGCAGAGATGGCTGTCTAAGCCCA	351	
QY	62	AlAarqValThrIlelyMetglucysAsnProAsnglnValAsnGlySerArqAsnSer	81	
DB	352	GCCAGGCGACCATCATGAATGAGAGTCAACCTACTAGGTGAATGGTTCCAGAGACTCA	411	
QY	82	ProAspArqCysSerValAlalasegllyGlyMetValSerSerSerAspAsnValGly	101	
DB	412	CCTGTAGTGCACAGTGTGAACAAAGTGGGAAGTGTGGCGGCCGATCTGTGGGG	471	
QY	102	MetAsnYrGlySerYrMetGlnGluIyrsHisIleProProProAsnMetThrThrAsn	121	
DB	472	ATGAGCTACGGCAGCTACATGAGAGAGAAGCATGTGGCCGCTCCCAATATGACCAAT	531	
QY	122	GluArqArqValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgln	141	
DB	532	GAGCGCAAGTGAATCGTCCGACAGATCTCTACTGTGTGAGACAGACCATGTCCGACAG	591	
QY	142	TrpIleuGluTrpAlaValIleGlyGlnTrpGlyLeuProAspValAspIleleuIleuHeGln	161	
DB	592	TGGCTGGAGTGGCGGTAAAGAAATATGGCCCTCGATATGTGGAGCTTACTATTTTCAG	651	
QY	162	AsnIleAspGlyIyGlnIleuCysIyMetThrIlyAspAspPheGlnArqLeuThrPro	181	
DB	652	AATATCGATGGGAAGAGAGCTGTGCAGATATGACAAAGATGATTCACAGGCGCTCACGCCG	711	
QY	182	SerYrAsnAlaAspIleleuIleuSerHisIleuHisYrLeuArgGluThrProLeuPro	201	
DB	712	AGTACAAATGGCGACATTTCTCTCCACATCTCCACTACCTCAGAGAGCTCCCTCCCA	771	
QY	202	HisleuThrSerAspAspValAspIyrsAlaleuGlnAsnSerProArqIleuMetHisAla	221	
DB	772	CATCTGACTTCGATGACGCTGTATAGGCTTTACAAAACTCTCCACGGTATATGCATGCC	831	
QY	222	ArqAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValYrProGluAlaThr	241	
DB	832	AGAAACACAGGGGGGCGAGCTTTATTTTCCAAATFACTGCTGATATCCGAGAGTAGG	891	
QY	242	GlnArqIleThrThrArqProAspLeuProYrGlnGlnAlaArqArqSerAlaTrpThr	261	
DB	892	CNAAGAAATTCACACTAGGCCAGATTTACTCTATATGAGCCCTCCAGAGATACACCTGAGAC	951	
QY	262	SerHisSerHisProThr---GlnSerIyrsAlaThrGlnProSerSerSerThrAlaPro	280	
DB	952	GGCCACAGCCACCTCACCCCTTCAGTCCAAAGCTGCTCAGCCCATTCCTCTGCAGTGGCC	1011	
QY	281	LysThrGluAspGlnArqProGlnIleuAspProYrGlnIleleuGlyProThrSerSer	300	
DB	1012	AAATCTGAAGACAGCGCTCTCAGTATATCTTACAGATCTTACAGATCTTGGACGACAGTAGC	1071	
QY	301	ArgIleuAlaAsnProGlySerGlyGlnIleGlnIleuTrpGlnPheIleuIleuGluIleuLeu	320	

Db 1072 CGCCTTCTATCCAGGATGAGCCAGATCCAGCTGTGGAGCTTCTGCTCGAATCTCTG 1131  
Qy 321 SerAspSerSerAsnSerAsnGysIlleThrTrrpelucllythrasnglygluphelymet 340  
Db 1132 TCACACAGCTCCACCTCCACATGCATCACCCTGGGAGAGCAACCAACGGGAGTTCCAAATG 1191  
Qy 341 ThrAspProAspGluValAlaIargArgTrrpelucllygluArglySerLysProAsnMetAsn 360  
Db 1192 ACACACCCGAGCAGAGTGCGCTCGGCGTGGGGAGAGAGAGCAACCAACCATATGAC 1251  
Qy 361 TyrAspLysLeuSerArgAlaLeuArgTrrTrrTrrAspLysAsnIleMetThrLysVal 380  
Db 1252 TATGACAAGCTCACCGCCGCTCCGCTACTACTACGACAAACATCATGACCAAGGTG 1311  
Qy 381 HisProProGluSerSerMetTrrLysTrrProSerAspLeuProTrrMetSerSerTyr 400  
Db 1311 ----- 1311  
Qy 401 HisGlyLysArgTrrAlaTrrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420  
Db 1312 CACGGAGAGCGCTACGGCTCAAGTTGACTTCCACGGAGATGGCCAGGCCCTGACGCC 1371  
Qy 420 ----- 420  
Db 1372 CACCCCTCTGAGTCTCCCTGTACACAGTACCCTCCGACCTGCATGAGGCTCTCTAT 1431  
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440  
Db 1432 CACGCCACCCCGAGAGATGAACTTGTGTCTCCACCCCTCCGCTCCCGACGTACA 1491  
Qy 441 SerSerSerPhePheAlaAlaProAsnProTrrTrrAsnSerProThrGlyIleTyr 460  
Db 1492 TCTTCCAGTTCTTGTGTCTCCGACACCATGACTGAGATTCACGAGATGGGGGATCTAC 1551  
Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTrrTyr 478  
Db 1552 CCGAACACTAGCTCCACGACGACATATGCTCTCCTGACCTGGGACCTTACTAC 1605

RESULT 4  
AY065661  
LOCUS 1413 bp mRNA linear VRT 17-DEC-2001  
DEFINITION Gallus gallus erg isoform C-1-1 mRNA, complete cds; alternatively  
spliced.  
ACCESSION AY065661  
VERSION AY065661  
KEYWORDS AY065661.1 GI:17887440  
SOURCE Gallus gallus.  
ORGANISM Gallus gallus.  
REFERENCE 1 (bases 1 to 1413)  
AUTHORS Iwamoto,M., Higuchi,Y., Koyama,E., Enomoto-Iwamoto,M., Kurisu,K.,  
Yeh,H., Abrams,W.R., Rosenbloom,J. and Pacifici,M.  
TITLE Transcription factor ERG variants and functional diversification of  
chondrocytes during limb long bone development  
J. Cell Biol. 150 (1), 27-40 (2000)  
JOURNAL 20351415  
MEDLINE 10893254  
PUBMED 10893254  
REFERENCE 2 (bases 1 to 1413)  
AUTHORS Iwamoto,M., Higuchi,Y., Enomoto-Iwamoto,M., Kurisu,K., Koyama,E.,  
Yeh,H., Rosenbloom,J. and Pacifici,M.  
TITLE The role of ERG (ets related gene) in cartilage development  
Osteoarthritis. Cartil. 9 Suppl A, S41-S47 (2001)  
JOURNAL 21535378  
MEDLINE 11680687  
PUBMED 11680687  
REFERENCE 3 (bases 1 to 1413)  
AUTHORS Iwamoto,M., Higuchi,Y., Yeh,H. and Pacifici,M.  
TITLE Direct Submission  
JOURNAL Submitted (05-DEC-2001) Oral Anatomy and Developmental Biology,  
Osaka University Faculty of Dentistry, 1-8 Yamadaoka, Suita, Osaka  
565-0871, Japan  
FEATURES Location/Qualifiers

source 1..1413  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
37..1392  
CDS  
/note="alternatively spliced; transcription factor"  
/codon\_start=1  
/product="erg isoform C-1-1"  
/protein\_id="AA140889.1"  
/db\_xref="GI:17887441"  
/translation="MASTIKELASVSDOSLEPCAYGSPHLAKTEMTASSSEYGO  
T SKSPRYPODDWLSQPPARVITIKMECNVNOVNGSRNSRDVCSVAKGKRWSSDNVGM  
NYSYMEKEHIPPNTNERRVTPADPTLMSYDHRDMLNFAVKEVGLPDVLI  
ONTDEKELKMTKDDFORLPYSNADILSHLYLRSGAFETIPNTSVYEAQRT  
TRDPLPYEORARSSAMTSHSPOTKATOPSSPTPKEDORPOLDPOIIGPTSR  
LRA NPSSGOIQLOMFLLELSDSNSNCTTMEGTNGEFKATDPEVARKKERKSNMAY  
DKLSRALRYTYDKNIMTKVHKGRYAYFDHGLAOLQPPPESSMTKYPBDELYMS  
YHAPQKMFVAPHPALPVTSSEFFAAPNPYMSPTGVIYPTNRLPAHAPSHLGT  
Y"

BASE COUNT 424 a 370 c 311 g 308 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,82e-146 Length: 1413  
Score: 2262.50 Matches: 430  
Percent Similarity: 86.35% Conservative: 0  
Best local Similarity: 86.35% Mismatches: 1  
Query Match: 87.42% Indels: 67  
DB: Gaps: 3

US-09-902-772-4 (1-478) x AY065661 (1-1413)

Qy 1 MetAlaSerThrIleLysGluAlaLeuSerValYalSerGluAspGlnSerLeuPheGlu 20  
Db 37 ATGGCAAGCACTATTAAAGAGCACTTATCAGTGTGAGTGAAGACCACTCTGTTGAG 96  
Qy 21 CysAlaTrrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
Db 97 TGTGCTTACGAGATGGCCCACTGCAAGACAGAAATGACAGCTCTCTCCAGAGAA 156  
Qy 41 TrrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrrPleuSerGlnPro 60  
Db 157 TATGGCGAACAATCAAGAGTACGCCCGGCTCCACGACGAGCTGTTATCACACCCC 216  
Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
Db 217 CCGGCCAGATTACCATTAAGAGAGTGAACCAACAGAGTTAATGGGTCAAGGAT 276  
Qy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
Db 277 TCACCTGATGCTGCACGCGGCAAAAGAGGAAATGGTTACACTTCAGACAAATGTT 336  
Qy 101 GlyMetAsnTrrGlySerTrrMetGluGlnLysHisIleProProAsnMetThrThr 120  
Db 337 GGGATGACATATGGAACCTCATGGAAGAGAACAGATTTCCGCTCCAAATATGACAAAC 396  
Qy 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrPserThrAspHisValArg 140  
Db 397 AATGACAGAGAGTTATTTGTCGACGAGATCTTACGTTATGAGACACAGACATGTACGG 456  
Qy 141 GlnTrrPleuGlnTrrAlaValLysGluTrrGlyLeuProAspValAspIleLeuPhe 160  
Db 457 CAGTGGCTGGAGTGGCGAGAGAGATGATGCTTCCACAGCGTGCACATCTGTTTTC 516  
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
Db 517 CAGACATATGATGGAAAGAGTGTGTAAATGACCAAAATGACTTCCAGAGACTTCACG 576  
Qy 181 ProSerTrrAsnAlaAspIleLeuLeuSerHisLeuHisTrrLeuArgValTrrProLeu 200  
Db 577 CCGAGCTTATACGCAATATCTCTCTGTCACCTTACACTACTCAGAGAGAGA----- 630  
Qy 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220

Db	630	-----	630
QY	221	AlaArgasnThrGlyGlyAlaThrPheIlePheProasnThrSerValTyrProGluAla	240
Db	631	-----GGAGCCACTTTTATTTTCCAAATACATCATGTTTACCACAAACA	675
QY	241	ThrGlnArgIleThrThrArgProaspLeuProTyrGluGlnAlaArgSerIleArg	260
Db	676	ACGCAAGAAATACAAACAAGGCGCAATTACCTTATGAGCAACGAGAGATCAGCGTG	735
QY	261	ThrSerHisSerHisProThrGlnSerIysAlaThrGlnProSerSerThrValPro	280
Db	736	ACGATTCACAGCCATCCCACTACGTCAAAAGCTACCACATCATCTTCAACAGTGGCC	795
QY	281	LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer	300
Db	796	AAAAAGAAAGACACACGCTCTCAGTTAGATCCTTATACGATTCCTTGAGCCGACGAGC	855
QY	301	ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPglInPheLeuLeuGlnLeu	320
Db	856	CGCTCTGGAAATCCAGGAGAGTGCGGAGATACAGCTATGGCACTTCTACTGAGCTCTG	915
QY	321	SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyTyrAsnGlyGlnPheIysMet	340
Db	916	TCGGACAGCTCCAACTCCAACTGCATACCTCGGGGACACAAATGGGAGTCAAGATG	975
QY	341	ThrAspProAspGluValAlaArgArgTrpGlyGluAlaGlySerIysProAsnMetAsn	360
Db	976	ACAGACCTCGATGAAGTGCTCGGCGTGGGGAGAGAGAAAGCAACCTAACATGAAAC	1035
QY	361	TyrAspIysLeuSerArgAlaLeuArgTyrTyrTyrAspIysAsnIleMetThrIysVal	380
Db	1036	TATGACAAACTCAGCCGCTGCACCTTCGTACTACTATGACAAATAATATGACTAAAGTT	1095
QY	380	-----	380
Db	1096	CATGGTAACGCTATGCCTACAAATTTATTTCCAGGAAATCGCTCAGCCCTCCAGCT	1155
QY	381	HisProProGluSerSerMetTyrIysTyrProSerAspLeuProTyrMetSerSerTyr	400
Db	1156	CACCTCCAGATCATCTCATGTATACAAATACCATCAGACCTCCCTACATGAGTTCCTAC	1215
QY	401	HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro	420
Db	1215	-----	1215
QY	421	HisAlaHisProGlnLysMetAsnPheValAlaProHisProIleAlaLeuProValThr	440
Db	1216	CATGCACACCCCGACAGAAATGATGACTTTAGCTCCCACTCCCGCTGGTCCCGGTAAAC	1275
QY	441	SerSerSerPhePheAlaAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleLeuTyr	460
Db	1276	TCATCCACACTTTTGTGCTGCCCTPATCATACGAGANTTCACCAACTGGAGGCACTACAC	1335
QY	461	ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr	478
Db	1336	CCCAATACAGAGCTGCGCAGCTGCATATATGCTTCCCATCTTGGCAGCCTACTACAC	1389
RESULT 5			
E31253			
LOCUS	E31253	1447 bp	DNA linear
DEFINITION	Protein having cell calcifying inhibitory activity and gene		
ACCESSION	E31253		
VERSION	E31253.1	GI:13025685	
KEYWORDS	JP 199075871-A/1.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1447)		
AUTHORS	Hiroyasu, I., Yoshinobu, H., Marjio, P., Joel, R. and Helena, E.		
TITLE	Protein having cell calcifying inhibitory activity and gene		
	encoding the same		

JOURNAL	Patent: JP 1999075871-A 1 23-MAR-1999. CHUGAI PHARMACEUT CO LTD, HE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA									
COMMENT	OS Unidentified PN JP 1999075871-A/1 PD 23-MAR-1999 PE 23-MAR-1998 JP 1998166076 PR 18-JUN-1997 US 08/878177, 20-JUN-1997 US 60/050297 PI HIOYASU IMAOTO, YOSHINOBU HIGUCHI, MARIUJO PASHIFIKI, PI JOEL ROZENBROOM, PI HELENA E PC C12N15/09, A61K48/00, C07K14/465, C07K16/18, C12Q1/68//A61K38/00 PC C12E21/02, PC C12P21/08, C12N15/00, A61K37/02 CC Strandedness: Single; CC Topology: Linear; FH Key FT source 1. 1447 Location/Qualifiers FT Location/Qualifiers 1. 1447 Location/Qualifiers FT Location/Qualifiers 1. 1447 Location/Qualifiers									
BASE COUNT	440 a 374 c 317 g 316 t									
ORIGIN										
Alignment Scores:										
Pred. No.:	1. 87e-146 Length: 1447									
Score:	2262.50 Matches: 430									
Percent Similarity:	86.35% Conservative: 0									
Best Local Similarity:	86.35% Mismatches: 1									
Query Match:	87.42% Indels: 67									
DB:	Gaps: 3									
US-09-902-772-4 (1-478) x E31253 (1-1447)										
QY	1 MetLaserThrilleySgluaLeuSerValSerGluAspGlnSerLeuPheGlu 20									
DB	63 ATGGCAAGCACTATTAAAGGAAGCATTTATCAGTGGAGGAAGACCACTCTGTTGGAG 122									
QY	21 CysAlaIArgIysSerProHisLeuAlaIySthGluMetThAlaSerSerSerGlu 40									
DB	123 TGTCCTTACGGATTCGCCCACTTGGCAAGACAGAAATGACAGCTCTCTTCCAGTGA 182									
QY	41 TyGlyGlnThSerIysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60									
DB	183 TATGGCAAAATCAAAAGATGAGACCCGGCGTCCCGACAGAGACTGGTTATCAGAGCC 242									
QY	61 ProAlaArgValThrIleIysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80									
DB	243 CCGCGCCAGAGTTACCTTAAGATGAGGTAAACCCAAACCAAGTTAAAGGCTCAAGGAAT 302									
QY	81 SerProAspAspCysSerValAlaIySgluIySgluMetValSerSerSerAspAsnVal 100									
DB	303 TCACCTGATATCAGTCAGCGCGGCAAAAAGAGGAAAATGGTTAGGACGTTACAGCAAACTT 362									
QY	101 GlyMetAsnIySgluSerIySgluMetGluGlnIySthHisIleProProProAsnMetThrThr 120									
DB	363 GGGATGACATATGGAAGCTACATGGAAGAGGAAGCATATTCGGCTCCAAATATGACAAAC 422									
QY	121 AsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140									
DB	423 AATGAACGAAGAGTATTGTCGCAGCAGATCTCCTACGTTATGGAGCACAGACATGTACGG 482									
QY	141 GlnTrpLeuGlnTrpAlaValIySgluIySgluIySgluLeuProAspValAspIleLeuLeuPhe 160									
DB	483 CAGGCGCTGAGGCGGCACTGAAGAGTGTGCTTCCAGACGTGACATCTTGTGTC 542									
QY	161 GlnAsnIleAspIlySgluLeuCysIySgluMetThrIySAspAspPheGlnArgLeuThr 180									
DB	543 CAGAACCTTATGAGGAAGAAGTTGTGTAAATAGACCAAGATGACTTCCAGAGCTACG 602									
QY	181 ProSerTyrAsnAlaIspIleLeuLeuSerHisIleuHisTyrLeuArgGluThrProLeu 200									

```

Db      603 CCGAGCTATACGAGATATCTCTGACACCTACACTACCTCAGAGAGAGA----- 656
Qy      201 ProHisLeuThrSerAspValAspLysAlaLeuInsSerProAlaGluMetHis 220
Db      656 ----- 656
Qy      221 AlaArgAsnThrGlyAlaThrPheProAsnThrSerValTyrProGluAla 240
Db      657 -----GGAGCCACTTTATTTTCCAAATACATAGTTACCCGAAAGCA 701
Qy      241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnAlaAlaArgSerAlaTyr 260
Db      702 ACGCAAGAAATACAAACAGCCAGATTACCTTATGAGCAAGCAGAGATCAGCGTGG 761
Qy      261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db      762 ACGAGTACACGCCATCCACTCAGTCAAAAGCTAACCCATCATCTTCCAAACAGTCCC 821
Qy      281 LysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db      822 AAAACAGACAGCCAGCGCTCCCTCAGTTAGATCCTTATCAGATCTTGGACCGACGACG 881
Qy      301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuGluLeu 320
Db      882 CGTCTGCAATCCAGGAGAGTGGGCAATACAGCTATGACGAGTCTTCTGAGACTTCTG 941
Qy      321 SerAspSerSerAsnSerAsnCysIleThrThrProGlnGlyThrAsnGlyLupHelyMet 340
Db      942 TCGACAGCTCCAACTCACTCACTGATCACCCTGGGAGGAGCAAAATGGGAGTTCAAGATG 1001
Qy      341 ThrAspProAspGluValAlaArgArgTyrGlyLupArgLysSerLysProAsnMetAsn 360
Db      1002 ACAGACCTGATGAAAGTGGCTCGGCGTTGGGAGAGAGGAAACCAACCTAACATGATGAC 1061
Qy      361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db      1062 TATGACAAACTCAGCCCTGCGACTTACTACTATGACAAATAATATTATGACTAAAGTT 1121
Qy      380 ----- 380
Db      1122 CATGGTAAACGCTATGCCCTCAAAATTTGATTTCCAGGAAATCGCTCAGGCCCTCAGCCT 1181
Qy      381 HisProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db      1182 CACCCCTCAGATATCATGATGATACAAATACCAATACCACTCCCTTACATGAGTTCTTAC 1241
Qy      401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1241 ----- 1241
Qy      421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db      1242 CATGCAACCCCGAGAGATGAATCTTGTAGCTCCATCCCTCGCTTGGCCGCTAAC 1301
Qy      441 SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyLysIleTyr 460
Db      1302 TCATCCAGCTTTTGTGTCGCCCTAATCCATACGGAATTCACCAACTGGAGGATCTTAC 1361
Qy      461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db      1362 CCCAATACCAAGGCTGCGACGCTGCTCATATGCTTCCATCTTGGCACCTACTAC 1415

```

RESULT 6  
 LOCUS AB073079 2133 bp mRNA linear ROD 17-OCT-2001  
 DEFINITION Mus musculus Erg mRNA, mouse homolog of human ets-related gene ERG,  
 transcript variant 2, complete cds.  
 AB073079  
 ACCESSION AB073079.1 GI:16197542  
 VERSION  
 KEYWORDS Mus musculus cDNA to mRNA, clone:Erg-2.  
 SOURCE  
 ORGANISM

**REFERENCE**  
 1 Ozawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.  
 Mus musculus Erg mRNA  
 Published only in Database (2001)  
 2 (bases 1 to 2133)  
 Ozawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.  
 Direct Submission  
 Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suenho-chou, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
**FEATURES**  
 source  
 1. 2133  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 /clone="Erg-2"  
 1. 2133  
 /gene="Erg"  
 /cds="149..1540"  
 /note="homolog of human ets-related gene (ERG), transcript variant 2"  
 /codon\_start=1  
 /protein\_id="BAB69949.1"  
 /db\_xref="GI:16197543"  
 /translation="MIQVPPPAHNIKALSVSEDSLEFCAATGPHLAKTEMTASSSDVGSRTKSPRYPOODWLSQAPARVITIKKPNQSOVNSRNSPDGCVKGGKMGSPDVGVGSSYSEMERKHPBPNTNRVLYVPADPLMSDHRQLRMLNAAVREYGLDVALEQNDGKELCKMKDKDRLRPLSTNADLLSLHLRLRPLPLHLSDDVDA LQNSPRMLRANRNGGAATFEPNVSVEATQRTTRAAQSPSAVETEDQRLQLP YQIAGPTSSRIANPNSGQIQIWMQFLLELSPSSNCTITWGTGGERKMDPDVAR WGERKSRPNMNYDLRSALRYDKNTKVGKRYAVKFFPHGIAQALQHPPESSLYKPSDLPYMGSVYAHNPQKMFVPHRPALPVTSSEFPASNPNWNSPTGSIYNTRL PASHMPSHLGTYV"

**BASE COUNT**  
 587 a 579 c 531 g 436 t  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 5,34e-141 Length: 2133  
 Score: 2186.00 Matches: 413  
 Percent Similarity: 85.31% Conservative: 11  
 Best Local Similarity: 83.10% Mismatches: 11  
 Query Match: 84.47% Indels: 62  
 DB: 10 Gaps: 3  
 US-09-902-772-4 (1-478) x AB073079 (1-2133)  
 Qy 2 AlaSerThrIleLysGluAlaLeuSerValValSerGlnAspGlnSerLeuPheGluCys 21  
 Db 173 GCACCTCATATTTAGAGGCGCTTGTCACTGTGAGCGAGGACGATCTTGAAGTGT 232  
 Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41  
 Db 233 GCCCTACGGAAGCCACACCTGGCTAAAGACAGATGACCGCATCTTCCAGTACTAT 292  
 Qy 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPheLeuGlnPro 61  
 Db 293 GGCACAGATCCAGAGATGATCCAGAGTCCCTTACGAGGAGCTGCTGTCTCAAGCCCA 352  
 Qy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
 Db 353 GCGAGGCTCACCATCAAGATGAGATGCAACCTTGTGAGTGAATGTTCCAGCAACTCA 412  
 Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101  
 Db 413 CCTGATGATGCACTGTGAACAAAGTGGGAGAGTGTGTGGGACCCGAGATACTGTGGG 472

OY	102	MetAnrYrGLySerTyrMeGlnLulYrYshSIlleProProAsnMethrThrAsn	121
Db	473	ATGACACTGCGACGCTACATGGAGAGAACATGGCCGCTCCCAATATGACCAAAAT	532
OY	122	GUArGaVAlleValProAlaAspProThrLeuTrpSerThrAspHisValArgLn	141
Db	533	GAGCCCAAGTGAATGTCCTCCAGATCTCTACTGTGGAGACAGACATGTCCGACAG	592
OY	142	TrPLeuGluTrPAlaValLysGluTrYrGlyLeuProAspValAspIleLeuLeuPheGln	161
Db	553	TGGCTGGAGTGGCGGTGAAGAAATATGGCCCTCGATGTGGAGAGCTGTACTATTTAG	652
OY	162	AsnIleAspGlyLysGluLeuCysLysMethTrIlyAspAspPheGlnArgLeuThrPro	181
Db	653	AATATCGATGGGAAGAGACTGTGCAGAAATGACAAAGATGACTTCACGGCTCAGCGG	712
OY	182	SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro	201
Db	713	AGCTCAATATGGCGACATTCCTCTCCATCTCCACTCCCTCAGAGACATCCCTTCCA	772
OY	202	HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMethIsla	221
Db	773	CATCGACTTCGCGATGAGCTGTATAGGCTTTACAAACTCTCCACGGTTAATGATGCC	832
OY	222	ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr	241
Db	833	AGAAACACAGGGGGGACAGCTTTTATTTTCCAAATACTGATAGTATATCCGAACTACG	892
OY	242	GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaThrPthr	261
Db	893	CAAAGAAATTACACTAGGCCA-----	913
OY	262	SerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValProLys	281
Db	914	-----GCTGTCGACCATGCTCCCTGTGACAGGCCAA	946
OY	282	ThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArg	301
Db	947	ACTGAGACCCAGCGCTCTCAGTTGATCTTACACGATCTGGGACGACCAAGTAGCGCG	1006
OY	302	LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeuSer	321
Db	1007	CTTGCTATTCAGAGTAGTGGCCAGATCCAGCTGTGGAGTTCCTGCAACTCCTGCA	1066
OY	322	AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMethThr	341
Db	1067	GACACCTCCAACTCCAACTGTGATCACCTGGGAGGACCAACGGGAGTTCAAGATGCA	1126
OY	342	AspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnThr	361
Db	1127	GACCCGAGCAGAGTGTGCTGGCGCTGGGGGAGAGAGAACCAACCATGATGACTAT	1186
OY	362	AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMethThrLysValHis	381
Db	1187	GACAAAGCTACACCGCGCTCCGCTACTACACCAAAACATCATGACCAAGTG---	1243
OY	382	ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerTyrHis	401
Db	1244	-----CAC	1246
OY	402	GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro---	420
Db	1247	GGGAAGCGCTACGGCTTCAAGTTTGACTTCACAGGGAATGGCCAGGCCCTCAGCCCA	1306
OY	421	-----His	421
Db	1307	CCTCCTGAGTGTCCCTGTACAAGTACCCCTCGACACTGCATACATGGCTCCTATAC	1366
OY	422	AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer	441
Db	1367	GCCCAACCCCAAGAAATGAATCTTTGTGTGCTCCACCACTCCCGCTCTCCAGTACATCT	1426
OY	442	SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro	461

Db	1427	TCCAGTTTCTTTGGCTTCCCGACACCCATACGATTCACCGACATCCACCG	1486
Qy	462	AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr	478
Db	1487	AACACTAGGCTCCACGACGACATATGCGCTCTCACACGGGACCTACTAC	1537
RESULT 7			
HUMERG2			
LOCUS	HUMERG2	3166 bp ss-RNA	linear PRI 08-NOV-1994
DEFINITION	Human erg2 gene encoding erg2 protein, complete cds.		
ACCESSION	M17254		
VERSION	M17254.1	GI:182186	
KEYWORDS	erg 2 protein.		
SOURCE	Human cell line COLO 320, CDNA to mRNA, clone lambda 12.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 3166)		
AUTHORS	Rao,V.N., Papes,T.S. and Reddy,E.S.		
TITLE	erg, a human ets-related gene on chromosome 21: alternative		
	splicing, polyadenylation, and translation		
JOURNAL	Science	237 (4815), 635-639	(1987)
MEDLINE	87263429		
PUBMED	3299708		
FEATURES			
source	Location/Qualifiers		
	1..3166		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map_21q22.3"		
	257..1645		
gene	/gene="ERG"		
CDS	257..1645		
	/gene="ERG"		
	/note="erg 2 protein"		
	/codon_start=1		
	/protein_id="AAA52398.1"		
	/db_xref="GI:182187"		
	/db_xref="GDB:G00-119-884"		
	/translation="MIVPDPAAHIKEALSVESDSLFECAYGTPLHANTENTASS		
	SSDYGQTKMSRPVPOQDMLSPAPRATVTKKCNPSQVNSRNSPDECSVAKGKMGV		
	SPDYGMNYSMEKHPENMTNERVIVPADPLTMSDHYROMLEMAVKRYGLE		
	DVNLIFQNDIGKELCKMTKDDPQRLPSPYNAADILSHLYLRREPLPLTSDVDKRA		
	LQNSBPLHARNTDLPYPPRPSAWTGGHPTPSKAQSPSPVPTREDPRLDPRT		
	QILGTSSTRLANPSSGQIQLMQFLLELSDSSNSCITWECTNBEFKTDDVARRM		
	GERKSPKMNWTDKLSRALRYVDKNINIMTKVGRKAYKFDHGLAQALQHPPPSSLYL		
	KYKSPDLPYMGSYHAHPQKMFVAPHPALPVTSSFFAAPPYMNSPGTIGYPTRTLP		
	TSNHPSLGTYT"		
	291..298		
repeat_region	/note="8-bp repeat sequence"		
repeat_region	509..516		
polyA_signal	/note="8-bp repeat sequence"		
	3143..3149		
	/note="putative"		
BASE COUNT	915 a 770 c 768 g 713 t		
ORIGIN	227 bp upstream of PvuII site; chromosome 21.		
Alignment Scores:			
Pred. No.:	2,06e-140	Length:	3166
Score:	2180.50	Matches:	413
Percent Similarity:	84.74%	Conservative:	9
Best Local Similarity:	82.93%	Mismatches:	11
Query Match:	84.25%	Indels:	65
DB:	9	Gaps:	4
US-09-902-772-4 (1-478) x HUMERG2 (1-3166)			
Qy	2	AlaserThrIleTylsGluAlaLeuSerValValSerGluAaSPGInserLeuPheGluCys	21
Db	281	GACGCTCATATCAAGAAAGCCTTATCAGTTGTGAGTAGAGACAGTCGTTTGGAGGCT	340
Qy	22	AlaTyrGlySerProHisLeuAlaIleTylThrGluMetThrAlaSerSerSerGluTyr	41



Dd	341	GCTAAGGAGACGCCACACTGGCTTAAGACAGAGTACGACCGCTCTCTCCACGACAT	400
Qy	42	GLYlnthrSerLysMetSerProargValProGlnInAspTrpLeuSerGlnPro	61
Dd	401	GGACAGACTTCCAAAGATGATGAGCCACGCGCTCCCTCAGCAGGATGTGCTGCTCAACCCCA	460
Qy	62	ALAargValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer	81
Dd	461	GCCAGGGCTACCATCAAAATGGAAAGTAACCTTACCAGGTGAAAGGCTCAAGAACTCT	520
Qy	82	PROAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly	101
Dd	521	CCTGATGATGATCAGTGTGGCCAAAGCGGGAGAGTGTGGCAGCCACGACCTTGGG	580
Qy	102	METAasnTrpGlySerTrpMetGluLubLysHisIleProProProAsnMetThrAsn	121
Dd	581	ATGAACTACGCGAGCTACATGGAGAGAAGCAGCATCCACCCCAACATGACCAAGAAC	640
Qy	122	GLUArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln	141
Dd	641	GAGCCCAAGATTATGCTGCACAGCAATCCTACGCTATGGAGATCAGACCATGTGGCGCAG	700
Qy	142	TRPLeuGlnTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPheGln	161
Dd	701	TGGCTGGAGTGGCGCGCTAAGAAATATGCTTCCAGACGTCACATCTTGTTATTCCAG	760
Qy	162	ASNileaspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro	181
Dd	761	AACATCGATGGGAAGAACTGTGCAGAAATGACCAAGACAGCATTCACAGCTCACCCCC	820
Qy	182	SerTrpAsnAlaAspIleLeuLeuSerHisIleuHisTrpLeuArgGluThrProLeuPro	201
Dd	821	AGCTCAACAGCGCGACATCTTCTCCATCATCTCCACTCACTCAGAGAACATCTCTTCCA	880
Qy	202	HISLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla	221
Dd	881	CATTGTGACTTCAAGATGATGTGATTAACCTTCAAAACTCTCCAGGGTTAATGATGACT	940
Qy	222	ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTrpProGluAlaThr	241
Dd	941	AGAAACACA-----	949
Qy	242	GlnArgIleThrThrArgProAspLeuProTrpGlnAlaAlaArgArgSerAlaTrpThr	261
Dd	950	-----GATTATCCATATGAGCCCCCAGAGATCAGCTGGACC	988
Qy	262	SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro	280
Dd	989	GGTCAAGGACACCCACGCCAGCGAAGCTGTCTCAACATCTCTTCCACACTGGCC	1048
Qy	281	LysThrGlnAspGlnArgProGlnLeuAspProTrpGlnIleLeuGlyProThrSerSer	300
Dd	1049	AAACTGAAGACACAGCTCTCAGTTAAATCTTTCAGATTCTTGGACCAACAAAGTACG	1108
Qy	301	ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeu	320
Dd	1109	CGCCTTGAATTCACGACGATGGCCAGATCCAGCTTGGCAGTTCTCTCGAGAGCTCTG	1168
Qy	321	SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMet	340
Dd	1169	TCGGACACCTCCAACTCCAGCTGCATCCCTCGGGAAGCACCAACGGGAGATTCAAGATG	1228
Qy	341	ThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsn	360
Dd	1229	ACGATTCGCCGACGAGGTGGCCCCGCGCTGGGGAGAGCCGAAGACCAACCAACTGTGAC	1288
Qy	361	TyrAspLysLeuSerArgAlaLeuArgTrpTrpTrpAspLysAsnIleMetThrLysVal	380
Dd	1289	TACGATAGCTCAGCGCGCCCTCGTCTACTACTATGACCAAGACATCATGCAAGAGTCT	1348
Qy	381	HisProProGlnSerSerMetLysTrpLysTrpProSerAspLeuProTrpMetSerSerTrp	400
Dd	1348	-----	1348

OY		401	HISGLYSARGTYRALATYLTLySPheNSPheNHSIGLYILEAGlnALeUgInPro	420
Db		1349	CATGGGAAGCCTACGCGCTTACAAGTTGCATTCCACGGGATGCCACAGGCCCTCCACGCC	1408
OY		420	-----	420
Db		1409	CACCCCCGGAGTCATCTCTACAAAGTACCCTTCAGACCTCCCGTACATGGGTCTAT	1468
OY		421	HISAlAHISPRoGInLysMeTaSnPhEVALABroHISPrOPrOIALeUPrOVA1Thr	440
Db		1469	CAGCGCCACCCACAGAAGATGACTTGTTGGCGCCCAACCCCTCCAGCCCTCCCGTGACA	1528
OY		441	SerSerSerPhePheAALAProASnProTyrrTrpAsnSerProThrnGlyGlyIleTy	460
Db		1529	TCTTCCACTTTTTTTTGGTGGCCCCAACCCCATACGTAATTCACCACTGGGGGATATATAC	1588
OY		461	PROASnThrArgLEUPrOIALAHISMetProSerHisLeUGlyThrTYTYTyr	478
Db		1589	CCCAACACTTAGGCTGCCACAGCACGACATATGCTTCTCATCTGGGCACTTACTAC	1642
RESULT 8				
LOCUS	XLAJ4125	1798 bp	mRNA	linear VRT 21-JAN-2000
DEFINITION	Xenopus laevis erg gene (erg_A).			
ACCESSION	AJ224125			
VERSION	AJ224125.1 GI:5420045			
KEYWORDS	ERG gene; transcription factor.			
SOURCE	Xenopus laevis.			
ORGANISM	Xenopus laevis			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
TITLE	Xenopodinae; Xenopus.			
JOURNAL	1 (bases 1 to 1798)			
MEDLINE	Baltzinger,M., Mager-Heckel,A.M. and Remy,P.			
PUBMED	Xl erg; expression pattern and overexpression during development			
REFERENCE	Dev. Dyn. 216 (4-5), 420-433 (1999)			
AUTHORS	20099678			
TITLE	Dev. Dyn. 216 (4-5), 420-433 (1999)			
JOURNAL	Submitted (24-FEB-1998) Baltzinger M., UPR9005, MMDC, Centre			
MEDLINE	National de la Recherche Scientifique, 15, RUE Rene Descartes,			
PUBMED	67084, FRANCE			
FEATURES	Location/Ouallifiers			
source	1..1798			
	/organism="Xenopus laevis"			
	/db_xref="taxon:8355"			
	/clone="lambda ZapIt-erg A"			
	/clone_lib="lambda ZapIt"			
	/dev_stage="stage 30 embryo"			
	321..1778			
	/gene="erg"			
	321..1778			
	/gene="erg"			
	321..1778			
	/gene="erg"			
	/function="transcription factor"			
	/codon_start=1			
	/protein_id="CABA6566.1"			
	/db_xref="GI:5420046"			
	/db_xref="SPRREML:09W700"			
	/translation="MOMARIPTYEDSNALSVSIEDSLFECTYTPHLITENTMA			
	SSSDSGTSKMSPRVPOODLSOPFSRVTIKMCSNFSQVNGSRSSPDOSTGRSM			
	GGGSDGNMGVSYIEERKLSPMTNTERRVIVPADPTLMSTDHWQMDEMAIKEYG			
	LIPDVVLIFONIDCKELKMTKEDFHTLPSTYNADILSHLHYLRETPLPHLTSDVD			
	KALONSRLIMARTGASFTPESSVYOANORIPRODLSYEPSRSAMTHNPAP			
	SKAOSPSTVPKTEDRPOLDPYOIIGPTSSLRNPSGOIOLOMFLELSDSNMN			
	CIEMEGNEEFKKMPDEVARNGEGRSKNNMYDKLSRLRYDDKNITMKYHGKY			
	AKTKDFGICIAOLQHPPESTIKYIBSELPYMSSYRAHNPQKMFVAHPHPALVTSSS			
	FFAAPNVNMVSPGSTIYPNTRLPASHMSHSLGYT"			
BASE COUNT	561 a 429 c 358 g 450 t			
ORIGIN				



CDS

63..1430  
/gene="VESPL4"  
/note="similar to human erg1"  
/codon\_start=1  
/product="vascular endothelial cell specific protein 14"  
/protein\_id="BAB62744.1"  
/db\_xref="gi:15128489"  
/translation="MASTIKEALSVSKDOSLFRCAYGTPLAKTEMTASSSDYQGT  
SKMSPRVQODMLSQPARVITKMECPNSOYNGSRNPDECSYTKGKMGSPITVGH  
SYSTMEKHMPPNMTNERRVIVPADPLMTSDHVRQLEMAVKEKGLADVLE  
ONIDKELCKMTKDFOFLTPSVNADILSLHAIKRETPLEHLSDDVDKALQMSPL  
MHANPDLPEPPEPSTWTSHTSHSPSKAQPSPSTVPTEDORPOLDPIYOLIGPS  
SLRANPSGOLOLWFLLESDSSNSCTWEGTEGFKMTDEDEVARMGKESKP  
NMNDKLSRAIRYYDKNIMTKVGRYAVKFDHGLAOLQPHPPSSSLTKYPSDL  
YMSYTHHPKMTNVAHPHPALPTVSSFFATPMPNPSPTGIGTIPNTRLPASHMPSH  
Lgity."

BASE COUNT 510 a 522 c 426 g 350 t

ORIGIN

Alignment Scores:

Pred. No.: 1.67e-140 Length: 1808  
Score: 2177.50 Matches: 412  
Percent Similarity: 84.37% Conservative: 9  
Best Local Similarity: 82.57% Mismatches: 13  
Query Match: 84.14% Indels: 65  
DB: 10 Gaps: 4

US-09-902-772-4 (1-478) x AB031088 (1-1808)

QY 1 MetAlaSerThrIleTylGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
DB 63 ATGGCCAGCACTATTAGAGAGCCTTGCTGCTTGAGCAGACAGCAGCTACTATTGAG 122  
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
DB 123 TGTGCTACGAGACGACACACCTGCTAGACAGAAATGACCGCATCTCTCCAGTGAC 182  
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPheSerGluPro 60  
DB 183 TATGCGCAGACATCCAGATAGTCCAGAGTCCACACAGAGAGCTGCTCTCAGGCC 242  
QY 61 ProAlaArgValThrIleLysMetGluLysAsnProAsnGlnValAsnGlySerArgAsn 80  
DB 243 CCAGCCAGGCTACACATCAAGATGAGTGCACACCTGACCGAGTGAAGGTTCCAGAGAC 302  
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
DB 303 TCACCTGACGATGACAGTGTGACCAAGGTGGGAAGATGGTGGGACGCTGATACGTG 362  
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120  
DB 363 GGAATGAGCTACGCGACGCTACATGAGAGACACATGCGGCCCCCAATATACACACG 422  
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140  
DB 423 AATGACGACAGAGTATGCTCCCTGCAGATCTCTCTGTGAGACACAGACAGTCCGG 482  
QY 141 GlnTyrPheGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
DB 483 CAGTGGCTGAGTGGGAGTGAAGAAATACGCGCTCTAGACGTGACGTCTTATATT 542  
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
DB 543 CAGATATATGACGGAGAGAGCTGTGCAAGATGCAAAAGTGCACCTCCACAGGCTACT 602  
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200  
DB 603 CCGAGGTACAAATGCGACATCTCTCTCACAATCTCCACATACCTCAGAGAGACCCCTT 662  
QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220  
DB 663 CCACATCTGACATCCGAGACGTTGATAGGCTTACAAAACCTCCACAGGTTAATCAT 722

QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240  
DB 723 GCTAGAAACACA----- 734  
QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnAlaAlaArgSerAlaTrp 260  
DB 735 -----GATTACTTATAGAGCTCCACAGAGATCAACACTGG 770  
QY 261 ThrSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerThrVal 279  
DB 771 ACCGGCCACACCCACCCACCCCTAGTCCAAAGCTGCCACCATCTCCCTCCACACTG 830  
QY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299  
DB 831 CCCAAACTGAAGACACACCGCTCTCAGTTAATCATCTTACCAAGATCTGGACCCACCACT 890  
QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeu 319  
DB 891 AGTGGCTTGCTAATCCAGTAGTGGCCAGATCCAGCTGTGGAGTTCTCTGGAACCTC 950  
QY 320 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLupheLys 339  
DB 951 CTGTCTGACACCTCCACTCCCACTGCATCAGCTGGGAAGGACACCAAGGGGAGTTCAAG 1010  
QY 340 MethThrAspProAspGluValAlaArgArgTrpGlyGluArgGlySerLysProAsnMet 359  
DB 1011 ATGACAGACCCGAGTAGAGTGGCTGGGGGGGGGCGGAGAGCAAGCAACCCCAACACTG 1070  
QY 360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379  
DB 1071 AACATATGACAAACACAGCGCTGCGCTCCGCTACAGTACAGCAAAAACATATGACCAAG 1130  
QY 380 ValHisProGluSerSerMetCylLysTyrProSerAspLeuProTyrMetSerSer 399  
DB 1131 GTG----- 1133  
QY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419  
DB 1134 ---CACGGGAAGCGCTATGCTTACCAAGTTTACTTCCACGGGATTTGCCAGGCCCTGCAG 1190  
QY 420 Pro----- 420  
DB 1191 CCCCATCCCCCGAGTGTGCTCCGTGACAAAGTACCCCTCCAGCTGCCGTATAGTGGCTCC 1250  
QY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439  
DB 1251 TATCACACACCCACCCGAAAGATGAACTTGTGCTCCGCCCTCCGGCCCTCCAGTTC 1310  
QY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyGlyTle 459  
DB 1311 ACATCTTCCAGTTCTTGTGCTACCCCGAACCACATCTGGAATTTGCCGACTGGGGGCATC 1370  
QY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478  
DB 1371 TACCGAACAACCTAGAGTCCCGACGACGACATATGCGCTCCACCTGCGGAGACTACTAC 1427

RESULT 10  
AB073078 2131 bp mRNA linear ROD 17-OCT-2001  
LOCUS AB073078  
DEFINITION Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG,  
transcript variant 1, complete cds.  
ACCESSION AB073078  
VERSION AB073078.1 GI:16191716  
KEYWORDS  
SOURCE Mus musculus cDNA to mRNA, clone:Erg-1.  
ORGANISM Mus musculus  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and  
Sakaki, Y.  
TITLE Mus musculus Erg mRNA  
JOURNAL Published only in Database (2001)

REFERENCE	2 (bases 1 to 2131)
AUTHORS	Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (15-OCT-2001) Masahira Hattori, The Institute of Physical

	BASE COUNT	582 a	590 c	530 g	429 t
	ORIGIN				
 Alignment Scores:					
Pred. No.:	1,71e-138	Length:	2131		
Score:	2149.50	Matches:	408		
Percent Similarity:	84.14%	Conservative:	11		
Best Local Similarity:	81.93%	Mismatches:	14		
Query Match:	83.06%	Indels:	65		
DB:	10	Gaps:	4		
 US-09-902-772-4 (1-478) x AB073078 (1-2131)					
OY	2 AlaSerThrIleLeuGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys	21			
Db	172 GCACCTCATATTATTAAGGAGGCCCTTGTCAGATTGTGAGCGAGCACCACTACTTTGAGTGT	231			
OY	22 AlaIyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluIyr	41			
Db	232 GCCCTACGGAAAGCCACACCCTGGCTTAAGACAGAATGACCCGATCTCTTCAGTAGCAT	291			
OY	42 GlyAlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro	61			
Db	292 GGCACAGACATCCAAGATGAGTCCCAGATCCCTTAGCGAGGATGGCTGTCTCAAGCCCCA	351			
OY	62 AlaArgValThrIleLysMetGluCysAnpProAnsnGlnValaAsngLysArgAsnSer	81			
Db	352 GCCAGGGGTACACATCAAGATGAATGACACACCCTAGTACGTGATGGTTCCAGGAACACA	411			
OY	82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValcyl	101			
Db	412 CCTGATGAGTGCACATGTGACCAAAGGTGGGAAGTGGGAGCGACCCCGATACTGTGGGG	471			
OY	102 MetAnfIyrGlySerIyrMetGluGluLysHisIleProProProAsnMetThrThAsn	121			
Db	472 ATGAGCTAACGGCACCTACATGAGAGAGAGAGATGTGGCGCTCCCAATATCACCAAAAT	531			
OY	122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgLn	141			
Db	532 GAGGCGAGAGATGATGTCTCTGCAGATCTCACTGTGTGGACCAAGACCAATGTCCGACAG	591			

OY	142	trpleuGlutPrpAlaValLysGluYrgrGlyLeuProAspValAspLleuleuPheGln	161
Db	592	TGGCTGGAGTGGGGCGGTGAAGAATAATGGCCCTTCGATGGAGCTCTTACTATTTCAG	651
OY	162	AsnLleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheGlnArgLeuThrPro	181
Db	652	AATATCGATGGGAAGAGCTGTGCAGAAATGACAAAGATGACTTCCAGGGCGTCACGGCG	711
OY	182	SerTyrAsnAlaAspLleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro	201
Db	712	AGCTCAATGGCCGACATCTCTTCACATCTCCACTCCTCAGAAAGATCCCTCCATCCA	771
OY	202	HisLauThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla	221
Db	772	CATCTGACTTCGCGATGACGTTATATAGGCTTTACAAACTCTCCACGGTTAATGCATGCC	831
OY	222	ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr	241
Db	832	AGAAACACA-----	840
OY	242	GlnArgLleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrpThr	261
Db	841	-----GATTTACTTATGAGCTCCCAAGGAGTATAGCTGGACC	879
OY	262	SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro	280
Db	880	GGCCACAGCCACCTCAGCCCTCAGTCCAAAGCTGCACAGCATCTCCCTCTGCAGTGC	939
OY	281	LysThrGlnAspGlnArgProGlnLeuAspProTyrGlnLleuleuLysProThrSerSer	300
Db	940	AAACTGGAAGACACCGTCTCAGTTAGTCCCTTACCAATCTCTGGACCGACCACTAGC	999
OY	301	ArgLeuAlaAsnProGlySerGlyGlnLleGlnLeuTrpGlnPheLeuGlnGluLeuLeu	320
Db	1000	CGCCTTGCTAAATCCAGTATGTGGCCAGATCCAGCTGTGGCAGTTCCTGCTCGAACTCTC	105
OY	321	SerAspSerSerAsnSerAsnCysLleThrTyrGlnGlyThrAsnGlyLeuPheLysMet	340
Db	1060	TCAGACACTCCAACTCCAACTGCATCACCCTGGGAAGCACCAACGGGGATGTCAGATG	1111
OY	341	ThrAspProAspGluValAlaArgArgTyrTrpGlyGlnArgLysSerLysProAsnMetAsn	360
Db	1120	ACAGACCCGAGAGAGGTGGCTGGCGCTGGGGGAGAGGAGAGGCAAGCCCAACATGAAAC	117
OY	361	TyrAspLysLeuSerArgAlaLeuAlaArgTyrTyrTyrAspLysAsnLleMetThrLysVal	380
Db	1180	TATGACAAAGCTCAGCGCGCGCTCCGCTACTGACGACAAATAATCATGACCAAGGTG	123
OY	381	HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr	400
Db	1239	-----	123
OY	401	HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro	420
Db	1240	CACGGGAAGCGGTAGCGCTACAGATTGACTTCCCAAGGATTTGCCAAGGCCCTCGACGCC	129
OY	420	-----	420
Db	1300	CACCCCTCTGAGTCTCCCTGTACAAAGTACCCCTCCGACCTGCACATACATGGGCTCTAT	135
OY	421	HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr	440
Db	1360	CACGGCCACCCCGAAGATGACTTTGTGTCTCCCAACCTCCGCTCTCCAGTCAACA	141
OY	441	SerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGlyLysLysLys	460
Db	1420	TCTTCACATTTCTTGTCTCCCGGAACCATACTGGAATTCCACCGCATGGGGCATCTAC	147
OY	461	ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr	478
Db	1480	CCGAACACTAGGCTCCACGACCATATGAGCCCTCAGCTTGGGGCACTTACATAC	1533

RESULT 11  
XLAJ4126 2012 bp mRNA linear VRT 21-JAN-2000  
LOCUS  
DEFINITION Xenopus laevis erg gene (erg\_E).  
ACCESSION AJ224126  
VERSION AJ224126.1 GI:5420047  
KEYWORDS ERG gene; transcription factor.  
SOURCE Xenopus laevis.  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
Xenopodinae; Xenopus.  
1 (bases 1 to 2012)  
Baltzinger, M., Mager-Heckel, A. M. and Remy, P.  
Xl erg: expression pattern and overexpression during development  
plead for a role in endometrial cell differentiation  
Dev. Dyn. 216 (4-5), 420-433 (1999)  
JOURNAL  
MEDLINE 10633861  
20099678  
REFERENCE  
AUTHORS 2 (bases 1 to 2012)  
Baltzinger, M.  
TITLE Direct Submission  
JOURNAL Submitted (24-FEB-1998) Baltzinger M., UPR9005, NMDCD, Centre  
National de la Recherche Scientifique, 15, RUE Rene Descartes,  
67084, FRANCE  
FEATURES  
SOURCE location/Qualifiers  
1..2012  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
/clone="lambda ZAPII-erg E"  
/clone\_lib="lambda ZAPII"  
/dev\_stage="stage 30 embryo"  
225..1595  
/gene="erg"  
225..1595  
/gene="erg"  
/function="transcription factor"  
/codon\_start=1  
/protein\_id="CABA6567.1"  
/db\_xref="GI:5420048"  
/db\_xref="SPRMBL:Q9W629"  
/translation="MIQTPDPSAHKEALSVSEDSLEFCTYGTPLHRTKTEWTASS  
SDYGVGTSKMSRPVODDMSOPSRVTIKMECPNSOVNSRSSPDSCISGSKMG  
GSDVGNNGYSYIEEKHSIPNMTNRRVTPADPTLMSDHYRDLKRNKRYGIP  
DYVYLFRONIDGKELCKMKEDPRILRTPSTNADILSHLHLRKGASFTFNSVTO  
DANORIPSRKDLSTERSRSAMTNHAPPSKASQSTVPTEDPRQDLDPYQILGPT  
SSRLANGSQIOLMOPLLELSDSNSNCTWEGTNGEFTMDPDEVARWGRKSK  
PNMYDKLSRLALRYDYDKNIMTKYHGKRYAKFPDGHIAQLQHPPESTMYKPSL  
PYMSYHAHPQKNFVAHPHPALPVTSSEFFAANAYWNSPTGSIYENTRLPSHMS  
HLGTRY"  
BASE COUNT 634 a 479 c 395 g 504 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.5e-129 Length: 2012  
Score: 2015.50 Matches: 383  
Percent Similarity: 81.38% Conservative: 19  
Best Local Similarity: 77.53% Mismatches: 23  
Query Match: 77.88% Indels: 69  
Gaps: 5  
US-09-902-772-4 (1-478) x XLAJ4126 (1-2012)

378 TCGAAATGAGCTCCACGCTGCTGCCTCAACAGACGTGCTCCACAGCTCCATCCAGGCTG 437  
Qy 65 ThrilleysMetGluCysAsnProbsnGlnValnsnglySerAArgAsnSerProAsp 84  
Db 438 ACCATCAAGATGAGATGACMACCCCTAGTCATGATGATCAAGAGAGTCCCGGATGAC 497  
Qy 85 CysSerValAlaIysGlyGlyAsMetValSerSerSerAspAsnValGlyMetAsnTyr 104  
Db 498 TGCAGTATAGGAAAGAT 557  
Qy 105 GlySerYrMetGluGlyGlyHisIleProProProAsnMetThrThrAsnGluArg 124  
Db 558 GCGACGATACATAG 617  
Qy 125 ValIleValProAlaAspProThrLeuTyrSerThrAspHisValArgGlnTyrLeu 144  
Db 618 GTGATTTAGCTGCTGACCTTCTGAGTGTGGAGCAGATCATGTGACAGCTGCTGAG 677  
Qy 145 TrrAlaValIysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAsp 164  
Db 678 TGGCAATAGAAAGATAGCGCTTCCAGATGTCATGCTGCTTCCAAACATGAC 737  
Qy 165 GlyIysGluLeuCysIysMetThrIysAspAspPheGlnArgLeuThrProSerTyrAsn 184  
Db 738 GCGAAGGATTTATGACAGATGACCAAGAGAGATTTCCATGACGACGCGCAAGCTACAT 797  
Qy 185 AlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuProHisLeuThr 204  
Db 798 GCTATATCTCTCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 839  
Qy 205 SerAspValAspIysAlaLeuGlnAsnSerProArgLeuMetHisAlaArgAsnThr 224  
Db 839 ----- 839  
Qy 225 GlyIleAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIle 244  
Db 840 ---GGTGCCTTTATTTTCCAAACATCGTTTATCAAGAGCAACCAAGATC 896  
Qy 245 ThrThrArgProAspLeuProTyrGluGlnAlaArgAsnSerAlaTrrThrSerHisSer 264  
Db 897 CCCAGCAGGCAAGATTTATCATATGATGACCCCAAGAGAGATCTGTCGCAAAATCATCT 956  
Qy 265 HisProThrGlnSerIysAlaThrGlnProSerSerSerThrValProIysThrGluAsp 284  
Db 957 GCACCA---CCATCAAAAGCATCTCAACCA---CTTACCAAGATTCCCAAAACGAGAGAC 1010  
Qy 285 GlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 304  
Db 1011 CCCAGGCCACAGCTAGATCCCTTATCAGATTTCTTGCCCAACACAGCCGACTTGCAAT 1070  
Qy 305 ProGlySerGlyGlnIleGlnLeuTrrPglNpheLeuLeuGluLeuSerAspSerSer 324  
Db 1071 CCAGGAGTGGGAGATTCAGCTCTGCAATTTCTTACTGCAATCTTGGAGATGTTCC 1130  
Qy 325 AsnSerAsnGlyIleThrTrpPgluglyThrAsnGlyGluPheIysMetThrAspProAsp 344  
Db 1131 AACTCCACATGCTACTTCTGGGAGAGCAACATGGAATTTAAGATGACCGATCGAT 1190  
Qy 345 GluValAlaArgArgTrrPglGluArgIysSerIysProAsnMetAsnTyrAspIysLeu 364  
Db 1191 GAAATGCGCAAGACTTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1250  
Qy 365 SerArgAlaLeuArgTrrTyrTyrAspIysAsnIleMetThrIysValHisProProGlu 384  
Db 1251 AGCCGTCACCTGCTTACTACTAGATGAAATTTATGATGATGATGATGATGATGATG 1298  
Qy 385 SerSerMetIysTyrTrrProSerAspLeuProTrrMetSerSerTrrHisGlyIysArg 404  
Db 1299 -----CATGGCAAGCGC 1310  
Qy 405 TyrAlaTrrIysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro----- 420  
Db 1311 TATGCTACAAATTTGATTTTTCATGAGGATGCTCAAGCTCTTCAACCTCATCTCCAGAA 1370

QY	421	-----HisAlaHisPro	424
Db	1371	TCAACTATGTACAGTATCCCTCAGAACTTCATATATGAGCTCATACATGACATCCCA	1430
QY	425	GlnlysetasnphreValAlaProHisProProAlaLeuProValThrSerSerSerPhe	444
Db	1431	CAGAGATGAACTTGTGGCCCCCAGCCACCTTACCTGTGACATGCTCAAGTTTC	1490
QY	445	PheAlaAlaProAsnProTyrTTPAsnSerProThrGlyGlyIleTyrProAsnThrArg	464
Db	1491	TTTGCACACCCCTATGATGATGCAATTCACCAACTGGAAGTATTTTCCAAATACCTCGG	1550
QY	465	LeuProAlaAlaHisMetProSerHisLeuGlyThrTyr	478
Db	1551	CTGCACGCTAGCCATATGCTCTTCATCTTGAGAACCTACTAT	1592
RESULT 12			
LOCUS	HUMERG11	3126 bp	mRNA linear PRI 27-Apr-1993
DEFINITION	Human erg protein (ets-related gene) mRNA, complete cds.		
ACCESSION	M21535.1	GI:182182	
VERSION	M21535.1	GI:182182	
KEYWORDS	erg protein.		
SEGMENT	1 of 2		
SOURCE	Human, cell line COLO 320, cDNA to mRNA, lambda-7.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 3126)		
AUTHORS	Reddy, E.S., Rao, V.N. and Papas, T.S.		
TITLE	The erg gene: a human gene related to the ets oncogene		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 84 (17), 6131-6135 (1987)		
MEDLINE	87317608		
PUBMED	3476934		
FEATURES			
source	Location/Qualifiers		
CDS	1..3126	/organism="Homo sapiens"	
	/db_xref="taxon:9606"		
	195..1286	/note="erg1 protein"	
	/codon_start=1		
	/protein_id="AAA35811.1"		
	/db_xref="GI:182182"		
	/translation="MWGSDPTVGMNTGSYEERKHPMPNTTNERVYVADPTLMST		
	DHYROWLEMAVKEYGIDVNIILFONIDGKELCKMTKDQRTSPSNADILLSHLY		
	LRPTPLHLSDVDKALONSPLRHARNDLPEPRRSAMHGHPPTQSAADS		
	PSVPTKEODRQOLDPVOLIGPTSSRLANGSGOIOIMOLLELSDSSMSSCITWEG		
	TNGEFKMTDDEYARRRGEKSKPPNNAYDKLSRLARTYTKNIMTYHGRRIAYKEDF		
	HGIAQALQHPPESSLIKYPSDLPMGSIYAHPOKMFVAHPHPALPVSSSFFAAPN		
	PYWNSTPTGTYPTSLPTSHMPSHLGTYT"		
BASE COUNT	928 a 732 c 725 g 741 t		
ORIGIN	1 bp upstream from EcoRI site.		
Alignment Scores:			
	Pred. No.:	6.87e-115	Length: 3126
	Score:	1808.50	Matches: 342
	Percent Similarity:	82.23%	Conservative: 5
	Best Local Similarity:	81.04%	Mismatches: 10
	Query Match:	69.88%	Indels: 65
	DB:	9	Gaps: 4
US-09-902-772-4 (1-478) x HUMERG11 (1-3126)			
QY	78	SerArgAsnSerProAspAspCysSerValAlaIleGlyGlyIleGlyMetValSerSerSer	97
Db	150	TCAAGGACTCTCTCTGTGATGACAGTGTGGCCAAAGCGGGAAGATGGTGGCAGCCCA	209
QY	98	AspAsnValGlyMetAsnTyrGlySerTyrMetGluGluIleProProProAsn	117
Db	210	GACACCGGTGGATGAACTACGAGCAGTACATGAGAGAGAACATGACCCCAAC	269
QY	118	MetThrThrAsnGluArgValIleValAlaProAlaAspProThrLeuTyrSerThrAsp	137
Db	270	ATGACCAAGAACGAGCCAGAGTATCTCCAGCAGATCTTACGCTATGAGTACAGAC	329
QY	138	HisValArgGlnTyrLeuGluTTPAlaValIleGlyGlyTyrGlyLeuProAspValAspIle	157
Db	330	CATGTGGCGGAGTGGCTGGAGTGGCGGTAAAGAAATATGCGCTCCAGACGTCACATC	389
QY	158	LeuLeuPheGlnAsnIleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheGln	177
Db	390	TTGTTATTCAGAACATCGATGGAGAACTGTGCAAGATGACCAAGACGACTTCAG	449
QY	178	ArgLeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGlu	197
Db	450	AGGCTCACCCCGACGTCACCAAGCGGACATCTTCTTCACATCTCCACTCCACGAGAG	509
QY	198	ThrProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArg	217
Db	510	ACTCTCTTCACATTTTGACTTCAGATGATGTTGATTAAGCTTACAAACTCTCCAGG	569
QY	218	LeuMetHisAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyr	237
Db	570	TTAATGATGCTGAAACACA	590
QY	238	ProGluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluAlaIleArg	257
Db	591	-----GATTTACCATATGAGCCCCCAGAGAGA	617
QY	258	SerAlaThrThrSerHisSerHisProThr-----GlnSerValAlaThrGlnProSerSer	276
Db	618	TCAAGCTGACCGGTGACGGCCACCCCGCCAGTCGAAACCTGCTCAACATCTCCT	677
QY	277	SerThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGly	296
Db	678	TCCACAGTGGCCCAAACTGAAAGACAGCGCTCAGTTAGATCCTTATGATGATCTTGG	737
QY	297	ProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeu	316
Db	738	CCAACAAGTGGCGCGCTTGCAGAAATCCAGGAGGCGAGTCCAGCTTGGCGATTCTC	797
QY	317	LeuGluLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGly	336
Db	798	CTGGAGCTCTGTGGGACAGCTCCAACTCCAGCTGATCACTCGGGAGGACACCAACGG	857
QY	337	GluPheLysMetThrAspProAspGluAlaIleArgArgTrpGlyGluArgLysSerLys	356
Db	858	GAGTTCAAGATGACGAGTATCCGACGAGAGTGGCCGGCGGAGAGCGAAGACCAAA	917
QY	357	ProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIle	376
Db	918	CCCAACATGAACTACGATTAAGTACAGCCGCGCTCCGTTACTACTATGACAAAGAACATC	977
QY	377	MetThrLysValHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyr	396
Db	978	ATGACCAAGGTC-----	989
QY	397	MetSerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGln	416
Db	990	-----CATGGGAAGCGCTACGCTTACAAAGTTGACATTCCACGAGGATGGCCAG	1037
QY	417	AlaLeuGlnPro-----	420
Db	1038	GCCCTCAGCCCGACCCCGGAGTCACTGTACAACTACCCCTCAGACCTCCGCTAC	1097
QY	421	-----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAla	436
Db	1098	ATGGGCTCTATACGCCCAACCCAGAAAGATTAATTTGTGGCGCCCACTCCAGCC	1157
QY	437	LeuProValThrSerSerSerPhePheAlaAlaProAsnProTyrTTPAsnSerProThr	456
Db	1158	CTCCCGGTGACATCTTCCAGTTTGTGGTGGCCCAACCCATATGGAATTCCACCACT	1217
QY	457	GlyGlyIleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThr	476

Db 1218 GGGGGTATATACCCACAGCTAGGCTCCACACGACATATGCTTCTCATCTGGGCACT 1277

QY 477 TTTTyr 478  
|||||

Db 1278 TACTAC 1283

RESULT 13  
DRE249590

LOCUS DRE249590 1436 bp mRNA linear VRT 03-MAR-2000

DEFINITION Danio rerio mRNA for Fil-1 protein (fil-1 gene).

ACCESSION AJ249590

VERSION AJ249590.1 GI:6006476

KEYWORDS fil-1 gene; Fil-1 protein.

SOURCE Danio rerio.

ORGANISM Danio rerio.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 1436)  
Brown, L.A., Rodaway, A.R., Schilling, T.F., Jowett, T., Ingham, P.W., Patient, R.K. and Sharrock, A.D.  
Insights into early vasculogenesis revealed by expression of the Ets-domain transcription factor Fil-1 in wild-type and mutant zebrafish embryos  
Mech. Dev. 90 (2), 237-252 (2000)

TITLE JOURNAL  
MEDLINE 20108585  
PUBMED 10640707

REFERENCE 2 (bases 1 to 1436)  
Sharrock, A.D.  
Direct Submission  
Submitted (29-SEP-1999) Sharrock, A.D., School of Biological Sciences, University of Manchester, Oxford Road, Manchester, M13, 9PT, UNITED KINGDOM

FEATURES  
source  
Location/Qualifiers  
1..1436  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/dev\_stage="20-24hr embryo"  
45..1400  
/gene="fil-1"  
45..1400  
/gene="fil-1"  
/function="transcription factor"  
/codon\_start=1  
/evidence=experimental  
/product="Fil-1 protein"  
/protein\_id="CAB56832.1"  
/db\_xref="GI:6006476"  
/translation="MDGTKEALSVSESDSLPEPPYAAAPLPKTDMTASGTODYGO  
THKINPIPOQEMINOPVRYNKRREYDHNGSRSPDSCVSGCKNKWVGTEASOMY  
TGYMDEKCAPPPMWTNERVYTPADPSLMSPDHYKQMLDAIKETGLDITPAHS  
TDGELECKMSKIDFLRLTYSYNTVEVLISHLNYLRESSISYNTPSHADSPLAARD  
DASYIDAVRRRTGWSNMHSGSPVVSQSKNPDPRPDPDYOLIGPSSPLAANG  
SGQIQLOMQLLELSDSANAGCTTWEETNEEFKMTDPEVARRMGEKSKPMNYDL  
SRALRYDYKNIMTKYGRKRAYKDFEHGIAOLAPRTESTYKYPSELYPYSYHA  
HQQKVNFSHPHPSMPTYSNFEFGPTTPYSSPPGGIYRIPSPVRIANSHVPSHLSY  
Y"

BASE COUNT 377 a 439 c 349 g 271 t

ORIGIN

Alignment Scores:

Pred. No.: 1 51e-93 Length: 1436

Score: 1491.50 Matches: 289

Percent Similarity: 68.13% Conservative: 53

Best Local Similarity: 57.57% Mismatches: 85

Query Match: 57.63% Indels: 75

DB: 5 Gaps: 7

US-09-902-772-4 (1-478) x DRE249590 (1-1436)

QY 1 MetaLaserThrilletySguAlaLeuSerValValSerGlunSpGlnSerLeuPheGlu 20  
|||||

Db 45 ATGACGCAACTATTAGGAGCGGCTGTCACTGTGTGATGTAAGACCACTCTCTGTTCGAG 104

QY 21 CysAlaTyr---GlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39  
|||  
::: ||| |||||:::|||||::: |||

Db 105 CTTCAATACGCGCGCTGCTCTCTTACCCAAAGACAGATGCTGGCCGGACAG 164

QY 40 GluTyrGluInThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 165 GATTACGGCCAGACGCAAAATATCAACCCATTCCCTTACGCAAGGTGATTATACAG 224

QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValaAsnGlySerArg 79  
||| ||||| |||  
::: ||||| |||

Db 225 ---CCGGTGGGGGGAACGTCAGACGAGAGAT-----GACCACATCAACGATTCAGA 275

QY 80 AsnSerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsn 99  
||||| ||||| ||||| |||  
||||| ||||| ||||| |||

Db 276 GAGTCCGCGGGTGAGCTCCACCGGGGAAATGCAATAAATGTGGGGCGGGAAGCG 335

QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGluTyrHisIleProProAsnMetThr 119  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 336 TCTCAGATGAACTACTACTGCTACATGATGAGAAATGTCTCCGCTCCAAACATACG 395

QY 120 ThrAsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisVal 139  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 396 ACCAATGAGAGAGATCATGCTCCCGCAGACCCGCTCTGTGGTCTCCGATACAGTG 455

QY 140 ArgGluTrpLeuGluTyrPalalaValLysGluTyrGlyLeuProAspValaLysLeu 159  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 456 CGGAGGTGTAAAGCTGGCCATTAAAGATACGCTTCAAGATGACACGCGCAT 515

QY 160 PheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeu 179  
||| ::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 516 TTTCACACGACAGATGAAAAAGACTCTGCAAGATGAGCAAAAGCACTTCGAGACTC 575

QY 180 ThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrPro 199  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 576 ACCAGCTTATTAAACACGAGGCTCTCTCATATCAATTAACCTCAGGAAAGT--- 632

QY 200 LeuProHisLeuThrSerAspAspValaLysLysAlaLeuGlnAsnSerProArgLeuMet 219

Db 632 ----- 632

QY 220 HisAlaArgAsnThrGlyAlaThrPheIleProAsnThrSerValTyrProGlu 239  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 633 -----ACGTCAATCATATCATACAAACGCGCATCTCAGCGTAC 671

QY 240 AlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAla 259  
::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 672 CAGTCCCTCGCTGCTGCAAAAAGACGATGCGTATGATGCTGACGAGAGACGGC 731

QY 260 TrpThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrVal 279  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 732 TGGTCGAACAACATGCACACCGGAAGGCTCCACAGTGTCTTCAGACGCTGCC 791

QY 280 ProLysThrGluAspGlnAAspProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 792 AAAAACCCGATACGCTGCTCTCCAGCCAGATCTTATAGATATTAGAACCCACACG 851

QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeu 319  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 852 AGTCGACTGCTTAACACGAGCTCAGTCAGATCCAGCTGAGCATGCTCTGTCGAGACTC 911

QY 320 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGluThrAsnGlyGluPheLys 339  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 912 CTGTCCACAGCGCCCAACGCGGCTGATACCTGGGAAGAACCAACGAGAGTTCACG 971

QY 340 MetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 359  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 972 ATGACGACCAACAGACAGGTGGCAGCGCTGGGGCGCAACGCAAGCAACCAACATG 1031

QY 360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1032 AACTACAGCAAAAGTGAACGCCGCTGCTACTACTACAGCAAAAGAACATCATGACCAA 1091

```

OY 380 ValHisProGluSerSerMetTyrTyrProSerAspLeuProTyrMetSerSer 399
Db 1092 GTG----- 1094
OY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspHisGlyIleAlaGlnAlaLeuGln 419
Db 1095 ---CACGGCAGAGCCCTACGCTACAGATTGACATTCCACGGCATCGCCAGCCCTGCA 1151
OY 420 Pro----- 420
Db 1152 CCCATCCACCGAGTCACCATGTACAGATCCCAAGCCAGCTGCCCTAGTGCCTAGC 1211
OY 421 ---HisAlaHisProGlnLysMetAspPheValAlaProHisProProAlaLeuProVal 439
Db 1212 TATACGCGCCACCGACGAGAGGTGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1271
OY 440 ThrSerSerSerPhePheAlaAlaProAspProTyrTyrPaspSerProThrGlyGlyLe 459
Db 1272 ACCTCTCCACACTCTTCTGCTGCCACCAACCCCATCTGGAGCTCCGCCACGAGAGTAT 1331
OY 460 TyrProAspThrArgLeuProAla-----AlaHisMetProSerHisLeuGlyThr 476
Db 1332 TACCCCATCCACGCTGCCACGCGCATGCGCAATTCACACGCTGCTTACACACTGTAGT 1391
OY 477 TyrTyr 478
Db 1392 TACTAC 1397

RESULT 14
AY029368 1359 bp mRNA linear PRI 09-MAY-2001
LOCUS Homo sapiens friend leukemia integration 1 transcription factor
DEFINITION (Fli1) mRNA, complete cds.
ACCESSION AY029368
VERSION AY029368.1 GI:14017402
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1359)
AUTHORS Upbi, B.T.S., Rainey, D.R. and Meredith, D.M.
JOURNAL Direct Submission
Submitted (09-APR-2001) Molecular Medicine Unit, Leeds University,
Beckett St, Leeds LS9 7TF, United Kingdom
FEATURES
source
1..1359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
1..1359
/gene="Fli1"
1..1359
/gene="Fli1"
/note="Ets transcription factor; involved in the
pathogenesis of Ewings sarcoma following translocation
with EWS gene on chromosome 22; may have oncogenic
potential when fused with EWS gene"
/codon_start=1
/product="friend leukemia integration 1 transcription
factor"
/protein_id="AAK50443.1"
/db_xref="GI:14017403"
/translat="MDGTIKELALSVSDSLPDSAYGAANLPRKADMTASGSPDYGO
PKINLPPOEININQPIRVKREYDHMGSRSPVDCSVSKSKYVGGSPPMY
NSYWDKNGPPPMPTNERVITVPADPTLMTQEHVROWLEMAIKEYSLMEIDSPFO
NMCKELCKMKKEDFLATLYNTVELLSLSTIKTEORQPDPTQILGPTSSRLAP
DPSTDYRKAWGNKNSGLKSPILGAGQFTSKTEORQPDPTQILGPTSSRLAP
GSGQILQMOFLLELSDANASCTTWTGTEGERKMDPDEVARWGERKSKPNMYDK
LSRALRYDKNITKYGKRAYKEDFHGIAQLPHPESSMYKPSDISYPSYH
AHQOKVNFVPPHPSMPVTSSTSPFGAASQVTSPTGITYPNPNVRHNPNTVHPSHLGS
Y"
BASE COUNT 359 a 422 c 333 g 245 t

```

```

ORIGIN
Alignment Scores:
Pred. No.: 6,35e-93 Length: 1359
Score: 1482.00 Matches: 291
Percent Similarity: 68.65% Conservative: 55
Best Local Similarity: 57.74% Mismatches: 80
Query Match: 57.26% Indels: 78
DB: 9 Gaps: 9

US-09-902-772-4 (1-478) x AY029368 (1-1359)
OY 1 MetaLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGln 20
Db 1 ATGACGAGGAGATATTAAAGAGGCTCTGCTGCTGAGCGACGACGATCCCTCTTGGAC 60
OY 21 CysAlaTyrGly----SerProHisLeuAlaLysThrGlnMetThrAlaSerSerSer 39
Db 61 TCAGCGTACGGAGCGGCGAGCCCATCTCCCAAGGCCACATGATGCTCGGGAGTCT 120
OY 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGln 59
Db 121 GACTACGGGAGAGCCGCCAAGATACACCCCTCCACACAGAGAGTGATCATACAG 180
OY 60 ProProAlaArgValThrIleLysMetGluCysAspProAsnGlnValAsnGlySerArg 79
Db 181 ---CCAGTAGGGCTCAACGCTCAACGCGGAGTAT-----GACCACATGAATGATCCAGG 231
OY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db 232 GAGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 291
OY 100 ValGlyMetAsnTyrGlySerTyrMetGluLysHis----IleProProProAsnMet 118
Db 292 AACCCCATGAACTACACAGCTATATGAGAGAGAGATGGCCCCCTCTCCCAACATG 351
OY 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTyrSerThrAspHis 138
Db 352 ACCACCAACGAGAGAGAGATGCTATGCTCCGCCACACCCACACTGTGGACACAGCAT 411
OY 139 ValArgGlnThrPheGlnGluTyrPheValValLysGluTyrGlyLeuProAspValAspIleLeu 158
Db 412 GTGAGCGAATGGCTGGAGTGGCCATTAAGAGATGACGTTGAGAGATGACACATCC 471
OY 159 LeuPheGlnAsnIleAspGlyLysGluLeuLysMetThrLysAspAspPheGlnArg 178
Db 472 TTTTCCAGAAATGATGATGCGAAGAGACTGTGTAATGAACAAGAGAGACTTCTCCGC 531
OY 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThr 198
Db 532 GCCACCAACCTCTACACACGAGAGAGTCTGTGTACACCTCACTCACTCAGGGAAGT 591
OY 199 ProLeuProHisLeuThrSerAspAspValAspIleLeuGlnAsnSerProArgLeu 218
Db 592 TCATCTG----- 597
OY 219 MethiAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
Db 598 -----CTGGCCTATATACACCTCCACACCC 624
OY 239 GluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSer 258
Db 625 GACCAATCCCTACAGATGAGTGTCAAGAGACCTTCTTATGATCAGTCACAGAGAGA 684
OY 259 AlaThrPheSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
Db 685 GCTTGGGGCAATTAATGAATTTCTGGGCTCAACAAAGTCTCCCTTGGAGGGGCGCAA 744
OY 278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297
Db 745 ACGATCAGTAAAGATACAGAGCAAGCGCCCGACAGATCCGATCAGATCCTGGGCGCG 804
OY 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleLeuLeuThrGlnPheLeuLeu 317

```



```

|||||
Db 805 ACCAGAGTCGCTAGCCACCTCGAAGGGGAGATCCAGCTGTGGCAATTCCTCTG 864
Qy 318 GtluLeuSerAspSerSerAsnSerAsnCysIleThrTyrPgluGlyThrAsnGlyGlu 337
Db 865 GACTGTCTCCACAGCCGACCCGACCTGTATCCCTGGAGGGGACCAACGGGGAG 924
Qy 338 PhylSerMetThrAspProAspGluValAlaArgTyrPgluGlyThrGlySerMetPro 357
Db 925 TTCAAAATGACGAGACCCGATGAGTGGCCAGCGGGGCGAAGCAACCAAGCC 984
Qy 358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377
Db 985 AACATGATATGACAGCAAGCTGACCGGGCCCTCGCTTATACATATATAAACAATTATG 1044
Qy 378 ThrLysValHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMet 397
Db 1045 ACCAAAGTG----- 1053
Qy 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417
Db 1054 -----CACGGCAAAAGATATGCTTACAAATTTGACTTCCAGCGCATTTGCCAGGCT 1104
Qy 418 LeuGlnPro----- 420
Db 1105 CTCGACGACACATCCGACCGAGTGTCCATGTACAGTACCTTCTGACATCTCTACATG 1164
Qy 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
Db 1165 CTTTCTTACCATGCCGCCAGCAAGAGGTGACATTGTCTCCCTCCATCATCTCCATG 1224
Qy 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTyrTrpAsnSerProThrGly 457
Db 1225 CCGTCTCCTTCCCTCCACCTTCTTGGAGCGCATCAACAATGAGACCTCCGCCAGGG 1284
Qy 458 GlyIleTyrProAsn-----ThrArgLeuProAlaHisMetProSerHisLeu 474
Db 1285 GGAATCTACCCCAACCCCAACGTCGCCCGCATCTTACACCAACGATGCTTACACTTA 1344
Qy 475 GlyThrTyrTyr 478
Db 1345 GGCAGCTACTTAC 1356

RESULT 15
BC010115 2394 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, clone MGC:19589 IMAGE:3635042, mRNA, complete cds.
DEFINITION BC010115
VERSION BC010115.1 GI:14603315
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2394)
Strausberg, R.
Direct Submission
Submitted (02-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNU)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,

```

```

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stanlipop, S., Thomas, P.J.,
Tjongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNU at: http://image.llnl.gov
Series: IRAL Plate: 27 Row: 9 Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 257353.
FEATURES
source
1. 2394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:19589 IMAGE:3635042"
/tissue_type="Lymph, Burkitt lymphoma"
/clone_lib="NIH-MGC.8"
/lab_host="DH10B-R"
/note="vector: POTB7"
126. 1484
/codon_start=1
/product="Unknown (protein for MGC:19589)"
/protein_id="AAH10115.1"
/db_xref="GI:14603316"
/translaton="MDGTIKREALSVSDQSIFSDAYGAAILPKADMTASGSPDYGO
PKINILPQDEWINDPVRVYKREYEDHNGSSESDVSFSCSLVGGSESPNMY
NSYMDKNGPPEPPMTNERVYLPADPTLMTGHEWLEMAIKYSLEIDTSPFO
NMGRKELCKMKNEKDFLATYLYTVEVLISLVSRLSLAYWTSTHTPOSSRLSYKE
DPSEYDVRGACAMNNMNSGLNKSPKPGAGOTISKMPORPOPPYDILPTSRILNP
GSGQIDLMQFLELSDSANASCTITEGNGEKMTDPDEVARMEERKSRPMKMTDK
LSRALRYVDKNTMTKRVKRVKFDPEHIALDLPHESSMYRVPADISYH
AHQKXNFVPPHPSMPVSSSFFGAASOYWTSPITGIYPNPVPRPHPTVPSHLS
Y"
BASE COUNT 669 a 609 c 574 g 542 t
ORIGIN
Alignment Scores:
Pred. No.: 1.27e-92 Length: 2394
Score: 1482.00 Matches: 291
Percent Similarity: 68.65% Conservative: 55
Best Local Similarity: 57.74% Mismatches: 80
Query Match: 57.26% Indels: 78
DB: 9 Gaps: 9
US-09-902-772-4 (1-478) x BC010115 (1-2394)
Qy 1 MetAlaSerThrIleLysGluAlaLeuSerValSerGluAspGlnSerLeuPheGlu 20
Db 126 ATGGACGGGACTATTTAAGAGGCTCTGCTGAGCAGACAGCAAGCTCCCTTTGAC 185
Qy 21 CysAlaTyrGly---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
Db 186 TCAGCGTACGAGGAGCGGAGCCATCTCCCAAGCGGACAGATGACTGCTGGGAGATCC 245
Qy 40 GluTyrGlnThrSerLysMetSerProArgValProGlnLysAspTrpLeuSerGln 59
Db 246 GACTACGGGAGCCGCCACAGATCAACCCCTCCACCAACAGCAGAGTGTGATCAATCAG 305
Qy 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db 306 ---CCAGTGAAGGCTCAACGTCGAAGCGGAGATAT-----GACCACATGAATGATCCAG 356
Qy 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db 357 GACTCTCCGCTGACTGACAGCGCTTACCAATGACAGAGCTGTGGCGGAGCGAGTCC 416
Qy 100 ValGlyMetAsnTyrLysSerTyrMetGluGluLysHis---IleProProProAsnMet 118
Db 417 AACCCCATGAACTACACAGCTATATGAGACAGAGAGATGGCCCTCTCCCAACATG 476
Qy 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuThrSerThrAspHis 138

```

|||||  
Db 477 ACCACGAGAGAGAGATGTCATGTCCTCCGACAGCCACACCTGTGCACAGAGCAT 536  
QY 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeu 158  
Db 537 GTGAGGCAATGGCTGGAGTGGCCATTAAGAGATGACGCTTGATGGAGATGACACATCC 596  
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178  
Db 597 TTTTCCAGAAATGATGCGAAGAACTGTGTAATGACAGAGAGAGACTTCTCCGCG 656  
QY 179 LeuThrProSerTyraAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThr 198  
Db 657 GCCACCACTCTCTACACACGGAAGTGTGTGCACACCTCAGTACCTCAGGGAAGT 716  
QY 199 ProLeuProHisLeuThrSerAspValAspLysAlaLeuGlnAsnSerProArgLeu 218  
Db 717 TCACTG----- 722  
QY 219 MethIAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238  
Db 723 -----CTGGCCTAATATACACCTCCACACC 749  
QY 239 GluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSer 258  
Db 750 GACCAATCTTCACGATGAGTCAAGAAGACCTTCTTATGACTCAGTCAGAGAGA 809  
QY 259 AlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277  
Db 810 GCTTGGGGCAATATACATGAATTCGGCCTCACAAGTCTCTCCCTTGGAGGGGCACAA 869  
QY 278 ThrValProLysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297  
Db 870 ACGATCAGTAAGATACAGCAAGCAAGCCCGCCAGCAGATCCGTATCAGATCCCTGGCCCG 929  
QY 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeu 317  
Db 930 ACCAGCAGTCGCTAGCCAACTGGAAAGGGGAGATCCAGCTGTGCAATTCCTCTG 989  
QY 318 GluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyIu 337  
Db 990 GAGCTGCTCCGACAGCGCAAGCAGCAGCTGTATCAGCTGGAGGGGACCAAGGGAG 1049  
QY 338 PheLysMetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysPro 357  
Db 1050 TTCAAATGACGAGCCCGATGAGTGCCAGGCGCTGGGCGAGCGGAAAGCAAGCCC 1109  
QY 358 AsnMetAsnTyraAspLysLeuSerArgAlaLeuArgTyrTyrTyraAspLysAsnIleMet 377  
Db 1110 AACATGAATACGACAGAGCTGAGCGGCGCTCGTATTACATGATATAAAACATTATG 1169  
QY 378 ThrLysValHisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMet 397  
Db 1170 ACCAAAGTG----- 1178  
QY 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417  
Db 1179 -----CACGGCAAAAGATATGCTTACAAATTGACTTCCACGGCATTGGCCAGGCT 1229  
QY 418 LeuGlnPro----- 420  
Db 1230 CTGAGCGCACAATCCGACGAGTGCATGACAAATACCTTCTGACATCTCCATCATG 1289  
QY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437  
Db 1290 CTTTCCATCAGATGCCACAGCAGAAAGTGAATTTCTCCCTCCCATCTCATG 1349  
QY 438 ProValThrSerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrGly 457  
Db 1350 CTTGTCACTTCTTCACTTCTTGTGAGCGCATACATATGACCTCCCGCAGGGG 1409  
QY 458 GlyIleTyrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474  
|||||

Db 1410 GGAATCTACCCCAACCCCAAGTCCCGCCATCTTAACACCCAGTGCTTCACTTA 1469  
QY 475 GlyThrTyrTyr 478  
Db 1470 GGCAGCTACTAC 1481

Search completed: July 28, 2003, 06:34:30  
Job time : 3830.11 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:05:39 ; Search time 306.661 Seconds

(without alignments)  
3510.246 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKKALSVSEDSQSLFE.....IYPTRLRPAHMSHCTYY 478

Scoring table:

BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPto.spool/US09902772/runtat\_23072003\_093656\_14868/app\_query.fasta.1.1294  
-DB=N\_Geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum2 -TRANS=human40.cdl  
-LIST=45 -DOCCALIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09902772.ecgn.1.1.364\_etunat\_23072003\_093656\_14868 -NCPu=6 -ICPu=3  
-NO\_MMAP -LARGESUBSTRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq.101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2436	94.1	1528	20	AAK26552	DNA encoding chick
2	2262.5	87.4	1447	20	AAK26551	DNA encoding chick
3	1482	57.3	2938	14	AAO50644	Human Hum-Fli-1 ge
4	1482	57.3	2937	24	ABK84139	Human cDNA differe
5	1464.5	56.6	2954	22	AAH02915	Human shear stress
6	1370	52.9	1890	14	AAO50662	Human Hum-Fli-1 ge
7	801	31.0	567	22	ABA48124	Human breast cell
8	801	31.0	567	22	ABA66003	Human foetal liver
9	801	31.0	567	22	ABA33090	Probe #11556 for g
10	801	31.0	567	22	AAK14426	Human brain expres
11	801	31.0	567	22	AAK40160	Human bone marrow
12	801	31.0	567	22	AAI20932	Probe #10865 for g
13	801	31.0	567	22	AAI46176	Probe #14862 used
14	801	31.0	567	22	AAI06643	Probe #6634 used t
15	801	31.0	567	24	ABSI4215	Human genome-deriv
16	716	27.7	473	22	ABA50193	Human foetal liver
17	716	27.7	473	22	ABA68128	Human foetal liver
18	716	27.7	473	22	ABA51512	Probe #13618 for g
19	716	27.7	473	22	AAK16512	Human brain expres
20	716	27.7	473	22	AAK42265	Human bone marrow
21	716	27.7	473	22	AAI48339	Probe #17025 used
22	716	27.7	473	22	AAI08696	Probe #8687 used t
23	654	25.3	533	21	AAO13442	Human secreted pro
24	565	21.8	549	24	ABK44114	CDNA #54 encoding
25	552	21.3	2025	23	ABLO5371	Drosophila melanog
26	503	19.4	420	22	ABA45056	Human breast cell
27	503	19.4	420	22	ABA55529	Human foetal liver
28	503	19.4	420	22	ABA52539	Probe #3705 for ge
29	503	19.4	420	22	AAK03760	Human brain expres
30	503	19.4	420	22	AAK29224	Human bone marrow
31	503	19.4	420	22	AAI35180	Probe #3866 used t
32	503	19.4	420	22	AAI03692	Probe #3683 used t
33	503	19.4	454	22	ABA42995	Human breast cell
34	503	19.4	454	22	ABA53411	Human foetal liver
35	503	19.4	454	22	ABA23187	Probe #1653 for ge
36	503	19.4	454	22	AAK01681	Human brain expres
37	503	19.4	454	22	AAK27129	Human bone marrow
38	503	19.4	454	22	AAI11718	Probe #1651 for ge
39	503	19.4	454	22	AAI33025	Probe #1711 used t
40	503	19.4	454	22	AAI01646	Probe #1637 used t
41	503	19.4	454	24	ABSO1685	Human genome-deriv
42	502.5	19.4	899	23	ABLI12097	Drosophila melanog
43	498	19.2	318	21	AAK21875	Human secreted pro
44	481	18.6	1752	22	AAK3672	CDNA encoding Rat
45	481	18.6	1752	24	AAI38753	Rat lambda73 CDNA

## ALIGNMENTS

RESULT 1  
AAK26552  
ID AAK26552 standard; DNA: 1528 BP.

XX AAK26552;

DT 14-JUN-1999 (first entry)

XX DNA encoding chicken c-ery protein.

DE Chicken; C-11 protein; cell calcification inhibiting activity;

KW cell calcification inhibiting agent; c-ery protein; arthritis deformans;

KW ossification; spinal column ligament; ss.

XX Gallus sp.  
OS  
XX  
FH Key  
CDS Location/Qualifiers  
63..1499

```

FT      /*tag= a
XX      JP11075871-A.
PN      23-MAR-1999.
XX      29-MAY-1998; 98JP-0166076.
XX      20-JUN-1997; 97US-0050297.
PR      18-JUN-1997; 97US-0878177.
XX      (CHUS ) CHUGAI PHARM CO LTD.
PA      (UYPE-) UNIV PENNSYLVANIA.
XX      WPI: 1999-257708/22.
DR      P-PSDB; AAY01521.
XX      An active protein for inhibiting cell calcification - useful for
PT      measuring the calcification of a cell, for diagnosing arthritis
PT      deformans or ossification of spinal column ligament
XX      Disclosure: Page 8-9; 15pp; Japanese.
XX      The present sequence encodes a chicken c-ery protein. The specification
CC      also describes a chicken C-11 protein (AAV01520) which has cell
CC      calcification inhibiting activity and a cell calcification inhibiting
CC      agent containing c-ery protein. The proteins are used for measuring the
CC      calcification of a cell, for diagnosing arthritis deformans or
XX      ossification of spinal column ligament.
XX      Sequence 1528 BP; 464 A; 395 C; 330 G; 339 T; 0 other.

Alignment Scores:
Pred. No.:      1,95e-192      Length:      1528
Score:          2436.00      Matches:      458
Percent Similarity: 91.97%      Conservative: 0
Best Local Similarity: 91.97%      Mismatches: 0
Query Match:    94.13%      Indels:      40
DB:             20      Gaps:        2

US-09-902-772-4 (1-478) x AAX26552 (1-1528)
QY      1 MetLaserThrIleLysGluLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      63 ATGGCAAGCACTAATTAAAGAGCAATTCAGTGTGAGTGAAGCAAGCTCTGTTGAG 122
QY      21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB      123 TGTCCTACGAGATGCCGCCACCTTGCAAGACAGAAATGACAGCTCTCTCCAGTGA 182
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
DB      183 TATGGCAACATCAAGAGATGAGCCCGCGCTTCCAGCAGGAGCTGTTATCATCACCCC 242
QY      61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB      243 CCGGCCAGAGTTACCATTAAGATGAGTGAACCAACCAAGGTTAATGGTCAAGGAAT 302
QY      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB      303 TCACCTGATGACTGACAGCGTGCAAAAGAGGAAATGGTTACCAATTCAGACAAATGTT 362
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProPropProAsnMetThr 120
DB      363 GGGATGAACATATGGAACCTACATGAGAGAGAGATATTCGCGCTCCAATATATACAAAC 422
QY      121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
DB      423 AATGAACGAGAGATTATGTGCCAGCAGATCTACGTTAAGAGACACAGACATGTACGG 482
QY      141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
DB      483 CAGTGGCTGGAGTGGGAGAGAGATGATGCTTCCAGACGTTGACATCTGTTGTTTC 542

```

```

QY      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB      543 CAGAACATTTGATGGGAAAGAGTTGTGTAATAATGACCAAAAGTACTTCAGAGAGCTCAGC 602
QY      181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB      603 CCGAGCTATTAACGAGATATCTCTCTGTCACACCTACATCACTCAAGAGAGACTCCCTT 662
QY      201 ProHisLeuThrSerAspAspValAspLysValLeuGlnAsnSerProArgLeuMetHis 220
DB      663 CCACATTTGACTTATGATGATGATGATTAAGGCTTACAAAACCTCCACGGTTAATGCAT 722
QY      221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
DB      723 GCTAGAAACACAGAGAGAGCCACTTTTATTTTCCAAATACATCAGTTTACCAGAGCA 782
QY      241 ThrGlnArgIleThrTrpArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 260
DB      783 ACGCAAAAGATTAACACAGAGCCAGATTACCTTATGAGCAGACGAGAGATCAGCTGG 842
QY      261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
DB      843 ACGAGTACAGCCATCCCACTCACTCAAAAGCTAACCCACATCATCTTACACAGTGCCTC 902
QY      281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
DB      903 AAAACAGAACAGCGCTCCATGATTAAGTTCATATGAGTTGTGAGACGACGACAGCAGC 962
QY      301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeu 320
DB      963 GCTCTTCCAAATCCAGGAGTGGGAGATACAGTATGAGCAAGTTCCTACTGAGCTTCTG 1022
QY      321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLysMet 340
DB      1023 TCGGACAGCTCCAACTCAACTGATCATCTGAGGAGGACCAAAATGGGAGTTCCAAATG 1082
QY      341 ThrAspProAspGluValAlaArgArgTrpGlyLysLysSerLysProAsnMetAsn 360
DB      1083 ACAGACCCTGATGAGTGGCTCGGCGTTGGGAGAGAGAGAAAGCAAACTAACATGAAAC 1142
QY      361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
DB      1143 TATGACAACTCAGCCGTGACCTGCTACTACTATGACAAATAATATATGACATAAAGTT 1202
QY      380 ----- 380
DB      1203 CATGCTAAAGCTATGCTTACAAATTGATTTCCACGGAATCGCTCAGGCCCTCCAGCT 1262
QY      381 HisProGlnLysSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
DB      1263 CACCCTCAGAAATCATCATGTCATGTAACATACCAATACCACTCCCTCATGAGATGCTTAC 1322
QY      401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
DB      1322 ----- 1322
QY      421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
DB      1323 CATGCAACCCCCAGAGATGAACCTTTGTAGCTTCCCATATCCCTGTGCCCCGCTGAAAC 1382
QY      441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 460
DB      1383 TCATCCAGCTTTTGTGTCGCCCTAATCCATCTGGAATTCACCACTGAGGACATCTAC 1442
QY      461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
DB      1443 CCCAATATCCAGGCTGCGAGCTGCTCATATGCTTCCATCTTGGACCTTACTAC 1496

```

RESULT 2

AAX26551 standard; DNA: 1447 BP.

XX

AC	AAI26551;
XX	
DT	14-JUN-1999 (first entry)
XX	
DE	DNA encoding chicken C-11 protein.
XX	
KW	Chicken: C-11 protein; cell calcification inhibiting activity;
KM	cell calcification inhibiting agent; c-ery protein; arthritis deformans;
KW	ossification; spinal column ligament; ss.
XX	
OS	Gallus sp.
XX	
FH	
FT	Key Location/Qualifiers
CDS	63..1418
FT	/tag= a
XX	
PN	JPI1075871-A.
XX	
PD	23-MAR-1999.
XX	
PF	29-MAY-1998; 96BP-0166076.
XX	
PR	20-JUN-1997; 97US-0050297.
PR	18-JUN-1997; 97US-0878177.
XX	
PA	(CHUS ) CHUGAI PHARM CO LTD.
PA	(TYPE-) UNIV PENNSYLVANIA.
XX	
DR	WPI; 1999-257708/22.
DR	P-PsDB; AAY01520.
XX	
PT	An active protein for inhibiting cell calcification - useful for
PT	measuring the calcification of a cell, for diagnosing arthritis
PS	deformans or ossification of spinal column ligament
XX	
PS	Disclosure; Page 7-8; 15pp; Japanese.
XX	
CC	The present sequence encodes a chicken C-11 protein which has cell
CC	calcification inhibiting activity. The specification also describes
CC	a cell calcification inhibiting agent containing c-ery protein
CC	(AAY01521). The proteins are used for measuring the calcification of a
CC	cell, for diagnosing arthritis deformans or ossification of spinal column
CC	ligament.
XX	
SQ	Sequence 1447 BP; 440 A; 374 C; 317 G; 316 T; 0 other;
	Alignment Scores:
	Pred. No.: 4,38e-178 Length: 1447
	Score: 2262.50 Matches: 430
	Percent Similarity: 86.35% Conservatve: 0
	Best Local Similarity: 86.35% Mismatches: 1
	Query Match: 87.42% Indels: 67
	DB: 20 Gaps: 3
OY	US-09-902-772-4 (1-478) x AAX26551 (1-1447)
Db	1 MetaLaserThrllelySGlAlaleuSerValValSergluNspInserleupheglu 20       63 ATGCAGACGACTATTAAGGACCATTTATCAGTGTGAGTGGAAGACCAGCTCTTTGGAG 122
OY	21 CyslaAtYrgylSerProHisLeualalysTrglumetrThAlaserSerSerglu 40       123 TGTCGCTCGGATGCCCCACCTTCCAAGAAGACGAATGACAGCCCTCTTCACGTGAA 182
OY	41 TyrGlglIntHserLySmetsSerProaragValProglnglasprHeuserSerglnpro 60       Db 183 TATGGGCAAACATCAAAAGATGGCCCCGCCGTTCCCAACAGAGATGGTTATCACAGCCC 242
OY	61 ProalaArqValThrlllelysmetglucySaSnProasnglValasnlySerAqasn 80       Db 243 CCGGCGAACAGTTCACATTAAAGATGAGTGTATACCCAAACAGAGTTATGGGTCAAGCAAT 302
OY	81 SerProaspPcySerValAlalalscglylysmetvalSerSerSeraspasnVal 1000

Dd	303	TCACCTGATGACTGCAGCGCTGGCAAAAGAGGAAAAATGGTTACGACGTTCAACAAATGTT	362
Qy	101	GIymeAsnTYrGLySerTYrMeGLuGLuLYshISileProProAsnMeThrThr	120
Dd	363	GGGATGAACCTATGGAAGCTACATGGAAGAAAGCATATTCGCCCTCCAAATATGACAAAC	422
Qy	121	AsnGLuArghrVAlIleValProAlaAspProThrLeuTPSerThrAspHisValArg	140
Dd	423	AATGAACACAGAGTTATTTGTGGCACACATCCTACGTTATGAGACACAGACCATATGTACGG	482
Qy	141	GLnTPleuGLuITrPAlaValLYsGLuYrGLyLeuProAspValAspIleuLeuPhe	160
Dd	483	CAGTGGCTGGAGTGGGCGACTGAGAGAGTATGGTCTTCCAGACGTGGACATCTTGTGTTTC	542
Qy	161	GLnAsnIleAspGLyLYsGLuLeuCYsLYsMeThrLYsAspAspPheGLnArgLeuThr	180
Dd	543	CAGAACATTTGATGGGAAAGAGTGTGTAAATGACCAAAAGATGACTCCAGAGACTCAG	602
Qy	181	ProSerTYrAsnAlaAspIleLeuLeuSerHisLeuHisTYrLeuArgIuThrProLeu	200
Dd	603	CCGAGCTTAAACGCGAGATATCCTCTGTCACACCTACACTCAGAGAGGA-----	656
Qy	201	ProHisLeuThrSerAspAspValAspLYsAlaLeuGLnAsnSerProArgLeuMeThIS	220
Dd	656	-----	656
Qy	221	AlaArgAsnThrGLyAlaThrPheIlePheProAsnThrSerValTYrProGLuAla	240
Dd	657	-----GGAGCCACTTTATTTTTCCAAAATACATCAGTTTACCACAAAGA	701
Qy	241	ThrGLnArgIleThrThrArgProAspLeuProTYrGLuGLnAlaArgIserAlaTrp	260
Dd	702	ACGCAAAAGATTAACAACAAGGCCAAGATTACCTTATGAGCAAGCAGAGATCACCGTG	761
Qy	261	ThrSerHisSerHisProThrGLnSerLYsAlaThrGLnProSerSerSerThrValPro	280
Dd	762	ACGAGTCAAGCCATCCACATCAGTCAAAAGCTTACCACACATCTTCAACAGTGGCC	821
Qy	281	LYsThrGLuAspGLnArgProGLnLeuAspProTYrGLnIleLeuGLyProThrSerSer	300
Dd	822	AAAAAGAAAGACACACGTCCTCAGTTAGATCCTTATCAGATTCTTGACCCGACGACGC	881
Qy	301	ArgLeuAlaAsnProGLySerGLyGLnIleGLnLeuTrpGLnPheLeuLeuGLuLeuLeu	320
Dd	882	CGTCTTGGAAATCCAGGGAGTGGGAGATACAGCTATGGCAgTTCCTACTGGAGGCTTCG	941
Qy	321	SerAspSerSerAsnSerAsnCYsIleThrTrpGLuGLyThrAsnGLyGLnPheLYsMet	340
Dd	942	TGGGACACCTCCAACTCCAACTGCATCCTGGGAGGCGACAAATGGGGAGTTCAAAGATG	1001
Qy	341	ThrAspProAspGLuValAlaArgArgTrpGLyGLnArgLYsSerLYsProAsnMetAsn	360
Dd	1002	ACAGACCTCGATGAAGTGGCTGGCGGTTGGGGAGAGAGAGAAAGCAAACTACATGAAAC	1061
Qy	361	TYrAspLYsLeuSerArgAlaLeuArgTYrTYrTYrAspLYsAsnIleMetThrLYsVal	380
Dd	1062	TATGCAAAACTCAGCGCGTGGCACTTGCTACTATGACAAAAAATATTATGACTAAAGTT	1121
Qy	380	-----	380
Dd	1122	CATGGTAAACGCTATGCTACAAATTTGATTTCCAGGAATCGCTCAGGCGCTCCAGCCT	1181
Qy	381	HisProProGLuSerSerMeTYrLYsTYrProSerAspLeuProTYrMetSerSerTYr	400
Dd	1182	CACCTCCAGAAATCATCTCATGTACAAATACCCATAGACGCTCCCTACATGAGTTCTTAC	1241
Qy	401	HisGLyLYsArgTYrAlaTYrLYsPheAspPheHisGLyIleAlaGLnAlaLeuGLnPro	420
Dd	1241	-----	1241
Qy	421	HisAlaHisProGLnLYsMetAsnPheValAlaProHisProProAlaLeuProValaThr	440

Db 1242 CATGCACACCCCGAGAGATGACTTTGTAGCTCCCACTCCCTGCTTTGCCCGTAACC 1301  
 Oy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLylleTyr 460  
 Db 1302 TCATCCAGACTTTTGTCTGCCCCCTTAATCAGTACTGAGATTCACACAGTGGAGCACTTAC 1361  
 Oy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478  
 Db 1362 CCCAATACACGAGTGGCCAGCTGTCATATGCTTCATGCTTGGACACTTACTAC 1415  
 RESULT 3  
 ID AA050644 standard; cDNA; 2938 BP.  
 AC AA050644;  
 XX 26-MAY-1994 (first entry)  
 DT  
 XX Human Hum-Fil-1 gene clone BM025.  
 XX  
 KW Chromosomal translocation; chimeric; chimaeric; Ewing sarcoma;  
 KW Ews gene; malignant melanoma; hum-fil-1;  
 KW primitive peripheral neuroectodermal tumour; human chromosome 11;  
 KW human chromosome 22; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 143..1501  
 FT /\*tag= a  
 FT /product= HUM-FIL-1  
 FT polyA\_signal 2908..2913  
 FT /\*tag= b  
 XX  
 XX MO9323549-A.  
 XX  
 XX 25-NOV-1993.  
 XX  
 XX 19-MAY-1993; 93WO-FR00494.  
 XX  
 XX 20-MAY-1992; 92FR-0006123.  
 XX  
 XX (CNRS ) CNRS CENT NAT RECH SCT.  
 XX  
 XX Aurias A, Delattre O, Desmaze C, Melot T, Peter M;  
 PI Plougastel B, Thomas G, Zucman J;  
 XX  
 DR WPI: 1993-386580/48.  
 DR P-PSDB: AAR44556.  
 XX  
 XX New nucleic acid of EWS gene and its hybrid(s) - contg. gene  
 PT sequence involved in chromosomal trans-location, also derived  
 PT mRNA, probes, fusion proteins etc., for diagnosis and treatment  
 PT of Ewing sarcoma and melanoma  
 XX  
 PS Disclosure; Fig 7; 123pp; French.  
 XX  
 CC The probe 11R1 was used to screen a human marrow cDNA library  
 CC (Clontech cat. # HL1058). The clone BM025 was identified and  
 CC sequenced. It represents the entire coding region together with  
 CC 5'- and 3'-UTRs of the Hum-Fil-1 gene.  
 XX  
 SQ Sequence 2938 BP; 847 A; 692 C; 675 G; 724 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 5.77e-113 Length: 2938  
 Score: 1482.00 Matches: 291  
 Percent Similarity: 68.65% Conservative: 55  
 Best Local Similarity: 57.74% Mismatches: 80  
 Query Match: 57.26% Indels: 78  
 DB: 14 Gaps: 9  
 US-09-902-772-4 (1-478) x AA050644 (1-2938)

Oy 1 MetaLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 143 ATGACGCGGACTATTAAAGAGGCTCTGTGGTGTGACGCGACGACCTCCCTTTTGAC 202  
 Oy 21 CysAlaTyrGly---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39  
 Db 203 TCACGCTAGCGAGGCGAGCCATCTCCCAAGGCCAGCATGCTGCTGGGAGAGCT 262  
 Oy 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59  
 Db 263 GACTACGCGGCCGCCAAGATCAACCCCTCCACACGACGAGGAGTGATCAATCAG 322  
 Oy 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79  
 Db 323 ---CCAGTGAAGGTCACAGTCACAGCGGAGAT-----GACCAATGATGATGCCAGG 373  
 Oy 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99  
 Db 374 GAGTCTCCGCTGAGCTCAGCGTTCAGCAATGACAGAGCTGTGGCGGAGCGAGCTCC 433  
 Oy 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---IleProProAsnMet 118  
 Db 434 AACCCCATGAACTAACACGCTATATGACGAGAGAGATGAGCCCTCTCCCAACATG 493  
 Oy 119 ThrThrAsnGluArgValAlleValProAlaAspProThrLeuTrpSerThrAspHis 138  
 Db 494 ACCACACAGGAGAGAGATGATGTCTCCGACACCCACACAGCTGTGACACAGAGCAT 553  
 Oy 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeu 158  
 Db 554 GTGAGGCAATGGCTGGAGTGGCCATTAAGAGATGATGCTTAAGAGATGACATATCC 613  
 Oy 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178  
 Db 614 TTTTCCAGACATGATGATGGCAGAGACTGTGTAATGACAGAGAGAGACTTCTCCGC 673  
 Oy 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThr 198  
 Db 674 GCCACACCCCTCATACACAGCAAGAGTGTGTGCACACCTGATCTCAGGAGAAAT 733  
 Oy 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218  
 Db 734 TCACCTG----- 739  
 Oy 219 MethHisAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238  
 Db 740 -----CTGGCTTATATACACCTCCACACC 766  
 Oy 239 GluAlaThrGlnArgIleThrArgProAspLeuProTyrGluGlnAlaArgArgSer 258  
 Db 767 GACCAATCTCCACGATGAGTGTCAAGAAGACCCCTTATATGACATCAGTCAAGAGAGA 826  
 Oy 259 AlaTrpThrSerHisSerHisProThrGlnSerLysAlaLhrGln---ProSerSerSer 277  
 Db 827 GCATGGGGCAATACATGAAATTCGTGGCTCAACAAAGATCTCCCTTGAGGAGGCCAA 886  
 Oy 278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297  
 Db 887 ACATATCGTAAAGATACAGAGCAAGGCCGCCACGACAGATCGTATGATGATCGTGGCCCG 946  
 Oy 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeu 317  
 Db 947 ACCAGCAGTGGCTAGCCCAACCTCGAAGGCGGACATCCAGCTGTGGCAATTCCTCTG 1006  
 Oy 318 GluLeuLeuSerAspSerSerAsnSerAsnCysIleIleThrTrpGluGlyThrAsnGlyGlu 337  
 Db 1007 GAGCTGCTCTCCACACGCGGACGCGAGCTGTATACCTGGGAGGAGGACCAACGCGGAG 1066  
 Oy 338 PheLysMetThrAspProAspGluValAlaArgArgTrpGlyLysLysSerLysPro 357  
 Db 1067 TTCAAAATGACGAGCCCGCATGAGTGCGCCAGGCGCTGGGCGGAGGAGGAGGAGGAGCC 1126

QY 358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrAspLysAsnIleMet 377  
 |||||  
 Db 1127 AACATGATATGACAGAGCTGAGCGGCCCTCGTTATATCTATGATTAATAACATATATG 1186  
 |||||  
 QY 378 ThrLysValHisProProGluSerSerMetLysTyrProSerAspLeuProTyrMet 397  
 |||||  
 Db 1187 ACCAAGAGT----- 1195  
 |||||  
 QY 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417  
 |||||  
 Db 1196 -----CACGGGAAAGATATGCTTACAAATTTGACTCCACGGCATTTGCCAGGCT 1246  
 |||||  
 QY 418 LeuGlnPro----- 420  
 |||||  
 Db 1247 CTGAGAGCAATCCGACGAGTCGTCATGTACAAAGTACCTCTTGACATCTCTACATG 1306  
 |||||  
 QY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeu 437  
 |||||  
 Db 1307 CCTTCCTACACATGCGCCACGACAGAGGTGAATCTTGCTCCCTCCCATCCATCCTCATG 1366  
 |||||  
 QY 438 ProValThrSerSerSerPhePheAlaProAsnProTyrTyrPasnSerProThrGly 457  
 |||||  
 Db 1367 CCTGTCACTCTCTCCAGCTTTTGGAGCCGATCAATATCTGAGCTCCGCCACGAGG 1426  
 |||||  
 QY 458 GlyIleTyrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474  
 |||||  
 Db 1427 GGAAATACACCCCAACCCAGCTGCCCGGCATCTTAACACCCAGTGCCTTACACTTA 1486  
 |||||  
 QY 475 GlyThrTyrTyr 478  
 |||||  
 Db 1487 GGCAGCTACTAC 1498  
 |||||  
 RESULT 4  
 ABR84139  
 ID ABR84139 standard; cDNA: 2957 BP.  
 AC ABR84139;  
 XX  
 DT 14-AUG-2002 (first entry)  
 XX  
 DE Human cDNA differentially expressed in granulocytic cells #710.  
 XX  
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
 KW viral infection; parasitic infection; protozoal infection;  
 KW fungal infection; sterile inflammatory disease; psoriasis;  
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
 KW adult respiratory distress syndrome; inflammatory bowel disease;  
 KW Crohn's disease; ulcerative colitis; periodontal disease;  
 KW granulocyte activation; chronic inflammation; allergy.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 FN WO200228999-A2.  
 PD 11-APR-2002.  
 XX  
 PE 03-OCT-2001; 2001WO-US30821.  
 XX  
 PR 03-OCT-2000; 2000US-237189P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 XX  
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
 XX  
 DR WPI; 2002-435328/46.  
 XX  
 PT Detecting granulocyte activation by detecting differential expression  
 PT of genes associated with granulocyte activation, which serves as  
 PT diagnostic markers that is useful for monitoring disease states and  
 PT drug toxicity -  
 XX  
 PS Claim 1; SEQ ID No 710; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation  
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
 CC DNA chip analysis as given in the specification, and comparing  
 CC the expression level to an expression level in an unactivated  
 CC GC, where differential expression of Gs is indicative of GCA.  
 CC Also included are modulating (M2) GA by contacting GC with an agent  
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)  
 CC for an agent capable of modulating GCA or an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease using the  
 CC gene expression profile; (3) detecting (M4) an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease, by detecting the  
 CC level of expression in a sample of the tissue of gene(s) from Gs, where  
 CC the level of expression of the gene is indicative of inflammation;  
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,  
 CC an allergic response in a subject, exposure of a subject to a pathogen  
 CC or sterile inflammatory disease, by contacting a tissue having  
 CC inflammation with an agent that modulates the expression of gene(s)  
 CC from Gs in the tissue. M1 is useful for detecting GCA. M2 is useful for  
 CC modulating GCA. M3 is useful for screening an agent capable of modulating  
 CC GCA preferably in an inflammation in a tissue; M4 is useful for  
 CC detecting an inflammation (especially chronic) in a tissue, an allergic  
 CC response in a subject, exposure of a subject to a pathogen or sterile  
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,  
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal  
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,  
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,  
 CC periodontal disease; also bacterial infection, viral infection,  
 CC parasitic infection, protozoal infection, fungal infection, and M5 is  
 CC useful for treating one of the above conditions. The present  
 CC sequence represents a gene differentially expressed in granulocytes.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX

Sequence 2957 BP; 855 A; 698 C; 680 G; 724 T; 0 other;

#### Alignment Scores:

Pred. No.:	5.82e-113	Length:	2957
Score:	1482.00	Matches:	291
Percent Similarity:	68.65%	Conservative:	55
Best Local Similarity:	57.74%	Mismatches:	80
Query Match:	57.26%	Indels:	78
DB:	24	Gaps:	9

us-09-902-772-4 (1-478) x ABR84139 (1-2957)

QY	1	MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu	20
Db	173	ATGGACGGGACTATTAAAGGAGCGCTCGGTGTGAGCAGCAGCCAGTCCCTTGGAC	232
QY	21	CysAlaTyrGly--SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer	39
Db	233	TCAAGCTACGAGGAGCGGAGCCCATCTCCCAAGGCCACATGACTGCTGGGAGTCT	292
QY	40	GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGln	59
Db	293	GACTACGGGAGGCCCAAGATCAACCCCTCCACACAGCAGAGGTGATCAATCAG	352
QY	60	ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg	79
Db	353	---CCAGTGAAGGCTCAACGTCAGCGGAGTAT-----GACCACATGATGATCCAGG	403
QY	80	AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn	99
Db	404	GAGTCTCGGTGACATCGACGCTTATGCAATATGACGACGACTGTGGGAGCGGAGTCC	463
QY	100	ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis--IleProProProAsnMet	118
Db	464	ACCCCATGAATACACAGCTATATGAGCAGAGAAATGCGCCCTCTCCCAACATG	523

```

QY 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHis 138
D 524 ACCACCAAGAGAGAGATGATGTCGCCGACACCCACACGCTGGACACAGAGCAT 583
QY 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeu 158
D 584 GTGAGGCAATGGCGGAGTGCATTAAGAGATGACGCTTGAAGATGACACATCC 643
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
D 644 TTTTTCGAGACATGATGGATGGCAAGAACTGTATAATGAAACAGAGAGACTTCTCCGC 703
QY 179 LeuThrProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThr 198
D 704 GCCACCAACCCCTACACACAGAGAGTGTCTGTGCACACCTCACTTACCTCAGGAGAGT 763
QY 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
D 764 TCACTG----- 769
QY 219 MethAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
D 770 -----CTGGCCTATAATACAACTCCACACC 796
QY 239 GluAlaThrGlnArgIleThrArgProAspLeuProTyrGluGlnAlaArgArgSer 258
D 797 GACCAATCTCAGCATGAGTGCATCAAGAGAACCTCTTATGACTCAGTCAGAGAGAGA 856
QY 259 AlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
D 857 GCTTGGGGCAATACATGAATCTCGCCCTCAACAAAGCTCCCTCCCTGGAGGGGCACAA 916
QY 278 ThrValProLysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297
D 917 ACCATCGTAAAGATACAGAGCAAGCGCCCGACGCAATCGTATCGATCTCGGCCCG 976
QY 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeu 317
D 977 ACCAGCAGTCGCTAGCCCAACCTCGAAGCGGCGATCCAGCTGTGCAATTCCTCCTCG 1036
QY 318 GluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrProGluGlyThrAsnGlyGlu 337
D 1037 GAGCTGCTCCGACAGCGCCAGCGCTGATCCCGGGGAGGAGGACCAAGCGGGAG 1096
QY 338 PheLysMetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysPro 357
D 1097 TTCAAAATGACGGACCCCGATGAGTGGCCAGCGCTGGGGCCGAGGAAAAGCAAGCCC 1156
QY 358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377
D 1157 AACATGAAATTACACAAAGTGAAGCGCGGCGCTCGTATATATATATAAACAATATATG 1216
QY 378 ThrLysValHisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMet 397
D 1217 ACCAAAGTG----- 1225
QY 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417
D 1226 -----CACGGCAAAAGATATGCTTACAAATTTGACTTCCAGGCAATGCCAGGCT 1276
QY 418 LeuGlnPro----- 420
D 1277 CTCGACGACATCCGACCGAGTGCATGTACAAATCCCTTCTGACATCTCCTACATG 1336
QY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
D 1337 CCTTCTACATGCCACAGCAGAGAGTAACTTGTCTCCCTCCCATCATCTCCATC 1396
QY 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpTrpAsnSerProThrGly 457
D 1397 CCGTGTACTTCCCTCCAGCTTCTTGGAGCGCATACAAATATCTGAGACTCCCCACGCGG 1456

```

```

QY 458 GlyIleTyrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474
D 1457 GGAATCTAACCCCAACCCCAAGAGTCCCGGACATCTTAACAGCCAGTGGCTTACACTTA 1516
QY 475 GlyThrTyrTyr 478
D 1517 GGCAGCTACTAC 1528
D 1517 GGCAGCTACTAC 1528
RESULT 5
AAH02915
ID AAH02915 standard; DNA; 2954 bp.
AC AAH02915;
XX
XX
XX 15-JUN-2001 (first entry)
DE Human shear stress-response coding sequence SEQ ID NO: 83.
KW Human; shear stress-response protein; vascular disease;
KW arteriosclerosis; ds.
XX
XX Homo sapiens.
XX OS
XX MO200125427-A1.
XX
XX 12-APR-2001.
XX
XX 02-OCT-2000; 2000MO-JP06840.
XX
XX 01-OCT-1999; 99JP-0280976.
XX
XX (KYOM) KYOMA HAKKO KOSYO KK.
XX (NOJI/) NOJIMA H.
XX Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
XX Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
XX WPI: 2001-266308/27.
XX P-PSDB: AAB90792.
XX
XX DNA sequences, proteins encoded by them and antihodies against them
XX useful in diagnosis and treatment of vascular disease caused by
XX arteriosclerosis -
XX
XX Claim 20: Page 462-466; 678bp; Japanese.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human shear stress response proteins. These are useful in the
XX diagnosis, treatment and screening of vascular diseases caused by
XX CC arteriosclerosis, including heart failure, post-PTCA restenosis and
XX CC hypertension.
XX
XX Sequence 2954 BP; 846 A; 698 C; 683 G; 727 T; 0 other;
SO
Alignment Scores:
Pred. No.: 1.64e-111 Length: 2954
Score: 1464.50 Matches: 290
Percent Similarity: 68.45% Conservative: 55
Best Local Similarity: 57.54% Mismatches: 80
Query Match: 56.59% Indels: 79
DB: 22 Gaps: 10
US-09-902-772-4 (1-478) x AAH02915 (1-2954)
QY 1 MetaLaserThrIleLysGluAlaLeuSerValSerGluAspGlnSerLeuPheGln 20
D 173 ATGAGCGGACTATTAAGAGGCTCTGCTGAGTGAAGCAGACAGTCCCTTGTGAC 232
QY 21 CysAlaTyrGly---SerProHisLeuAlaValThrGluMetThrAlaSerSerSer 39
D 233 TCAGCGTACGAGGAGCGGAGCCCATCTCCCAAGCGCAGATGACTGCTGGGAGATCCT 292
QY 40 GluTyrGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59

```



Db	233	GACTACGGGGACCCCAACAGATCATACCCCTCCACCAACAGACAGAGATGATCATACG	352
OY	60	ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg	79
Db	353	---CCAGTGGAGGCTACAGTCACGGGGAGTAT-----GACCACATGATGATCCAG	403
OY	80	AsnSerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsn	99
Db	404	GAGTCTCCGGTGGACTGTCAGCGCTTGACAAATGCACACAGACCTGGTGGCGAGCGAGTCC	463
OY	100	ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---IleProProProAsnMet	118
Db	464	AACCCATGAACCTACACAGCTATATGACGGAAGAAATGGCCCCCTCTCTCCACATCG	523
OY	119	ThrThrAsnGluArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHis	138
Db	524	ACCACACAGAGAGAGATCATCTCTCCCGCAGACCCACACCTGTGGACACAGAGACAT	583
OY	139	ValArgGlnTyrPLeuGluTyrPalAValLysGluTyrGlyLeuProAspValAspIleLeu	158
Db	584	GTCAGGCAATGTGGTGGAGTGGCCCTAAAGGAGTACAGCTTGATGAGATCCACATCC	643
OY	159	LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg	178
Db	644	TTTTTCCGAACATGATGATGGACAGGAACTGTATAAATGAACACAGAGAGACTTCTCC	703
OY	179	LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluThr	198
Db	704	GCCACACACCTCTACACACAGAAATGCTGTTGTACACACTGATTACTCAAGGAAAT	763
OY	199	ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu	218
Db	764	TCACGTG-----	769
OY	219	MethSAlaIaArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrPro	238
Db	770	-----CTGGCCCTAATAAACAACCTGCCACACC	796
OY	239	GluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaIaArgAspSer	258
Db	797	GACCAATCTCTACAGTATGAGTGTCAAAGAAGCCCTTCTTATGACTCACTCAGAAAGAGA	856
OY	259	AlaIleThrThrSerHisSerHisProThrGlnSerLysAlaIaThrGln---ProSerSerSer	277
Db	857	GCTTGGGGCAATACATGATATCTGCGCTCACACAAAAGTCCCTCCCTGGAGGGGCACAA	916
OY	278	ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro	297
Db	917	ACGATCACTAAGAAATACAGAGACAGACAGCCCGCCAGCCAGATCCCTATACAGATCTCGGGCCCG	976
OY	298	ThrSerSerArgLeuAlaAsnProGlySerGluGlnIleGlnLeuTrpGlnPheLeuLeu	317
Db	977	ACCACGATCTGCGTACCAACCTTGGAAACGGGGCGAGATCCACACTGGGCAATTCCTCCGG	1033
OY	318	GluLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGluTyrHisAsnGlyGlu	337
Db	1037	GAGCTGCTCTCGACAGCGCCACCCAGCTGTATCACCTGGGAGGGGACCAAGGGGAG	1099
OY	338	PheLysMetThrAspProAspGluValAlaIaArgArgTyrPheGluArgLysSerLysPro	357
Db	1097	TTCAAAATAGACGAGACCCCGATGAGTGGCCAGCGCCCTGGGGGAGGGGAAAGCAAGCCC	1155
OY	358	AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet	377
Db	1157	AACATGAATTTACGACAAAGCTGAGCGGGGGCTCCGTTATTAATATGATTAACAACTTAAG	121
OY	378	ThrLysValHisProGluLeuSerGermTyrLysTyrProSerAspLeuProTyrMet	397
Db	1217	ACCAAAAGTG-----	1222
OY	398	SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla	417

Db	1226	-----CACGGCCAAAGATATGCTTACAAATTTGCATTCCACAGGCATTCGCCAGGCT	1276
Oy	418	LeuGIaPro-----	420
Db	1277	CTCGAGCACCACATCCGACCGAGTCGTCCATGTACAAGTAGCCCTTGATCATCTCCATCATG	1336
Oy	421	-----HSAIahISPROGLnlyMetasnPhyeValAlAPROHISPROFolaIeu	437
Db	1337	CCTTCCCAACACTGGCCACACAGACGAAGAAGTGAACCTTTGTCCTCCCATCCATCCATCCCATG	1396
Oy	438	ProvalThurSerSerPheAlaAlAProAsnProTyrtPranserProThgLy	457
Db	1397	CCTGTCACTTCCTCCAGCTTCTTTGGAGCCGCATACAAATACGTGACCTCC---ACGGGG	1453
Oy	458	GlylleTyrtProAsn-----ThrArgleuProAlalAalAhISmetProSerHisIeu	474
Db	1454	GGAAATTCACCCCAACCCCAACGTCGCCCGCATGCTTAACACCCACGTGCTTCACACTTA	1513
Oy	475	GlyThrtTyrtYr	478
Db	1514	GGCAGCTACTAC	1525
<b>RESULT 6</b>			
AAQ50662	ID	AAQ50662 standard; DNA: 1890 BP.	
XX	AAQ50662;		
AC			
DT	26-MAY-1994	(first entry)	
XX			
DE		Human Hum-Fil-1 gene (genomic DNA).	
XX			
KW		chromosomal translocation; chimeric; chimaeirc; Ewing sarcoma;	
KV		Ews gene; malignant melanoma; hum-fil-1;	
KW		primitive peripheral neuroectodermal tumour; human chromosome 11;	
XX		human chromosome 22; ss.	
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	exon	1..39	
FT		/*tag= a	
FT		/number= 1	
FT		/codon_start= 22..24	
FT	Intron	40..99	
FT		/*tag= b'	
FT		/number= 1	
FT	exon	100..311	
FT		/*tag= c	
FT		/number= 2	
FT	Intron	312..371	
FT		/*tag= d	
FT		/number= 2	
FT	exon	372..526	
FT		/*tag= e	
FT		/number= 3	
FT	Intron	527..586	
FT		/*tag= f	
FT		/number= 3	
FT	exon	587..790	
FT		/*tag= g	
FT		/number= 4	
FT	Intron	791..850	
FT		/*tag= h	
FT		/number= 4	
FT	exon	851..916	
FT		/*tag= i	
FT		/number= 5	
FT	Intron	917..976	
FT		/*tag= j	
FT		/number= 5	
FT	exon	977..1042	
FT		/*tag= k	

FT /number= 6  
 FT Intron 1043..1102  
 FT /tag= 1  
 FT /number= 6  
 FT exon 1103..1162  
 FT /tag= m  
 FT /number= 7  
 FT Intron 1163..1222  
 FT /tag= n  
 FT /number= 7  
 FT exon 1223..1270  
 FT /tag= o  
 FT /number= 8  
 FT Intron 1271..1330  
 FT /tag= p  
 FT /number= 8  
 FT exon 1331..1890  
 FT /tag= q  
 FT /number= 9  
 FT /number= 9  
 PN W09323549-A.  
 PD 25-NOV-1993.  
 XX 19-MAY-1993; 93WO-FR00494.  
 XX 20-MAY-1992; 92FR-0006123.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 PI Aurias A, Delattre O, Desmaze C, Melot T, Peter M;  
 PI Plougastel B, Thomas G, Zucman J;  
 XX WPI: 1993-386580/48.  
 DR P-PSDB; AAR44556.  
 XX  
 PT New nucleic acid of EMS gene and its hybrid(s) - contg. gene  
 PT sequence involved in chromosomal trans-location, also derived  
 PT mRNA, probes, fusion proteins etc., for diagnosis and treatment  
 PT of Ewing sarcoma and melanoma  
 XX  
 PS Disclosure; Fig 7 and Fig 13; 123pp; French.  
 XX  
 CC The intron-exon junctions of the human Ews gene and the Hum-F11-1  
 CC gene have been sequenced (see AA050646 and AA050662, respectively).  
 CC The different fusion products which could be formed by fusing  
 CC exons from the two genes, as happens after specific chromosomal  
 CC translocations, can be predicted (see AA050671-050678). The sequences  
 CC at fusion junctions of other observed translocations are given in  
 CC AA050679-050683). NOTE: the 1890 residue sequence AA050662 does not  
 CC appear in the specification; it is a combination of the Hum-F11-1  
 CC cDNA sequence (Fig 7, AA050644) and the individual intron sequences  
 CC with their intron-exon junctions (Fig 13).  
 XX  
 SQ Sequence 1890 BP; 457 A; 561 C; 458 G; 414 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 6,12e-104 Length: 1890  
 Score: 1370.00 Matches: 294  
 Percent Similarity: 55.89% Conservative: 57  
 Best Local Similarity: 46.82% Mismatches: 83  
 Query Match: 52.94% Indels: 194  
 DB: 14 Gaps: 16  
 US-09-902-772-4 (1-478) x AA050662 (1-1890)  
 OY 5 ILILSGUALALeuSerValValSerGLASpGlnSerLeuPheGluCysAlaTyGly 24  
 DB 94 CTCAGAGAGCTCTGCTGCTGTGAGAGACAGACAGCTCTTGTGACTCAGGTAAGA 153  
 OY 25 ---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyGln 43  
 DB 154 GCGGACAGCCATCTCCCAAGGCCGACATGACTGCTCGGGAGAGTCTGACTACGGGACAG 213

OY 44 ThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnProAlaArg 63  
 DB 214 CCCACAAAGATCAACCCCTCCACACAGAGAGAGATCATCATGAG---CCAGTGAGG 270  
 OY 64 ValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySer-----78  
 DB 271 GTCACAGCTCAAGCGGAGTAT---GACCACATGATGATGCCAGTAAGTCACACAG 324  
 OY 79 -----ArgAsnSerProAsp 83  
 DB 325 GCGTGTGACAGATTGGGCTTGGGCTTTGGCCCTCTCCTCACTTAAAGAGTCTCCGGTG 384  
 OY 84 AspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsn 103  
 DB 385 GACTGACAGCGTTAGCAATGATGACAGACAGCTGGTGGCGGAGGCTCAACCCCATGAC 444  
 OY 104 TyrgLyserTyrmGluGluLysHis---IleProProProAsnMetThrThrAsnGlu 122  
 DB 445 TACACACAGCTATATGAG 504  
 OY 123 ArgArgValIleValPro-----128  
 DB 505 AGGAGAGTCAATCGTCCCGCAGAGTAATCGAAGAACAGGCTGCGGCGGCTCTTGGT 564  
 OY 129 -----AlaAspProThrLeuThrPserThrAspHisValArgGlnTrp 142  
 DB 565 AACACGCTCTTCCTCGCAGACCCACACTGTGACACAGAGAGATGAGAGCATGAG 624  
 OY 143 LeuGluTrpAlaValLysGluTyGlyLeuProAspValAspIleLeuPheGlnAsn 162  
 DB 625 CTGAGAGGGGCTTAAAGAGATAGCTTATGAGATGAGATGAGATGAGATGAGATGAG 684  
 OY 163 IleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheGlnArgLeuThrProSer 182  
 DB 685 ATGGATGGCAAGAGACGTGTAAATGAACAGAGAGAGATCTCTCGGCGCACACCCCTC 744  
 OY 183 TyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuProHis 202  
 DB 745 TACACACAG 804  
 OY 203 LeuThrSerAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAlaArg 222  
 DB 805 AGTACC-----CAG 813  
 OY 223 AsnThrGlyLysAlaThrPheIlePhe-----ProAsnThr 234  
 DB 814 GCGTGGGGTATTAACCTGTTATGTTTGGCTCTCAGGTTCACTGCTGAGCTATATATCA 873  
 OY 235 SerValTyPProGluAlaThrGlnArgIleThrThrArgPro-----248  
 DB 874 ACCTCCACACGACCAATCTCAGATGAGTGTCAAGAGAGTAGTGTGCTTTG 933  
 OY 249 -----AspLeuProTyGluGln 254  
 DB 934 TGCACCTAAATTAATGTACCCCTATTTGTTATGTTCAATTAAGCCCTTATGACTCA 993  
 OY 255 AlaArgArgSerAlaThrPThrSerHisSerHisProThrGlnSerLys-----270  
 DB 994 GTCAG 1053  
 OY 271 -----AlaThrGlnPro-----274  
 DB 1054 TTTTATAGTCTTTGGAGGCTCACTGATTTCTTCCCTCTGACACAGAGTCTCCGCTT 1113  
 OY 275 ---SerSerSerThrValProLysThrGlnAspGlnArgProGln-----288  
 DB 1114 GGAGGGGACAAACGATCACTAGTAAGTACAGACAAAGGCGCCAGGATACCTGCGCCA 1173  
 OY 289 -----LeuAspProTyGln 293  
 DB 1174 GGATATGTAATCTCTCTTTGAAGCAAAATTTCTTTTATTTCTTATGATCCGTATCAG 1233



Db 234 CCCCCGAGTACTCTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCTATAC 175  
 QY 422 AAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441  
 Db 174 GCCACCCACAGAAATGATGACTTTGTGGCGCCACCTCCAGCCCTCCCGGTGACATCT 115  
 QY 442 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleTyrPro 461  
 Db 114 TCCAGTTTTTTTGGCGCCCAAAACCCATACGATTCACCAACTGGGGGTATATACCC 55  
 QY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleTyrTyr 478  
 Db 54 AACACTAGGCTCCACCAAGCCATATGCTTCTCATCTGGGCATTTACTAC 4  
 RESULT 8  
 ABA66003/C  
 ID ABA66003 standard; DNA; 567 BP.  
 AC ABA66003;  
 AC ABA66003;  
 DT 01-FEB-2002 (first entry)  
 DE Human foetal liver single exon nucleic acid probe #14308.  
 XX  
 KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
 OS Homo sapiens.  
 PN WO200157277-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001WO-US000669.  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483447/52.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 FT analyzing gene expression in human fetal liver -  
 XX  
 PS Claim 4; SEQ ID NO 14308; 639bp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

US-09-902-772-4 (1-478) x ABA66003 (1-567)  
 QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGlnLeuLeuSer 321  
 Db 554 CTGACCTCATTAAGGAGAGTGGCCAGATCCAGCTTTGGAGTTCTCTCTGGAGCTCTGTGG 475  
 QY 332 AspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyGlnPheLysMetThr 341  
 Db 474 GACAGCTCAACTCCAGCTGATCATCCTGGGAAGACCAACGAGGAGTTCAAGATGACG 415  
 QY 342 AspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTyr 361  
 Db 414 GATCCCGAGGAGGTGGCCCGCGCTGGGGAAGACGGAACCAACCATGAACTAC 355  
 QY 362 AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 381  
 Db 354 GATAAGCTCAGCCGCGCCCTCCGTTACTACTATGCAAGAACATCATATCACAAGTCTC 298  
 QY 382 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401  
 Db 297 -----CAT 295  
 QY 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420  
 Db 294 GGGAAAGCGCTACGCTTACAGATTCGACTTCACAGGGATGCGCCAGGCCCTCCAGCCAC 235  
 QY 421 -----His 421  
 Db 234 CCCCCGAGTACTCTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCTCATAC 175  
 QY 422 AAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441  
 Db 174 GCCACCCACAGAAATGATGACTTTGTGGCGCCACCTCCAGCCCTCCCGGTGACATCT 115  
 QY 442 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleTyrPro 461  
 Db 114 TCCAGTTTTTTTGGCGCCCAAAACCCATACGATTCACCAACTGGGGGTATATACCC 55  
 QY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleTyrTyr 478  
 Db 54 AACACTAGGCTCCACCAAGCCATATGCTTCTCATCTGGGCATTTACTAC 4  
 RESULT 9  
 ABA33090/C  
 ID ABA33090 standard; DNA; 567 BP.  
 AC ABA33090;  
 AC ABA33090;  
 DT 23-JAN-2002 (first entry)  
 DE Probe #11556 for gene expression analysis in human heart cell sample.  
 XX  
 KM Human; gene expression; heart; microarray; vascular system; probe;  
 KM cardiovascular disease; hypertension; cardiac arrhythmia;  
 KM congenital heart disease; ss.  
 OS Homo sapiens.  
 PN WO200157274-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001WO-US000669.  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI: 2001-488899/53.  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts -  
 XX  
 PS Claim 4: SEQ ID NO 11556; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart. The  
 CC present sequence is one such probe. The probes may be used for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from the human heart via microarrays. By measuring gene expression, the  
 CC probes are useful for predicting, diagnosing, grading, staging,  
 CC monitoring and prognosing diseases of the human heart and vascular system  
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
 CC congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1.78e-57 Length: 567  
 Score: 801.00 Matches: 151  
 Percent Similarity: 78.17% Conservative: 3  
 Best Local Similarity: 76.65% Mismatches: 3  
 Query Match: 30.95% Indels: 40  
 DB: 22 Gaps: 2  
 US-09-902-772-4 (1-478) x ABA3090 (1-567)  
 QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 321  
 DB 534 CTGACCTCATTTAGGAGAGTGGCCAGATCCAGCTTGGAGTCTCTCTGGAGCTCTGCG 475  
 QY 322 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGluPheLysMetThr 341  
 DB 474 GACAGCTCCACCTCAGCTGCATCACCCTGGAGGACCAACGAGGAGTCAAGTACG 415  
 QY 342 AspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTrp 361  
 DB 414 GATCCCGAGGAGGAGGCGGCGTGGGAGAGGAGCAACCAACATGTAATAC 355  
 QY 362 AspLysLeuSerArgAlaLeuArgTrpTrpLysAspLysAsnIleMetThrLysValHis 381  
 DB 354 GATAAGCTCAGCCCGCCCTCCGTTACTACTATGCAAGAACATCATACCAAGGTC--- 298  
 QY 382 ProProGluSerSerMetLysTrpLysTrpSerAspLeuProGlyMetSerSerTrpHis 401  
 DB 297 -----CAT 295  
 QY 402 GlyLysArgTrpAlaTrpLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420  
 DB 294 GGGAGGCGCTACGGCTTACAGTTCGACTTCACAGGAGTCCAGGCCCTCCAGCCAC 235  
 QY 421 -----His 421  
 DB 234 CCCCCGGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGGTACATGGCTCTATAC 175  
 QY 422 AlaHisProGlnLysMetAsnPheValAlaLabronHisProProAlaLeuProValThrSer 441  
 DB 174 GCCACCCACAGAGATGATTTGTGTGGCCGCCCTCCAGCCCTCCCGGTGACATCT 115  
 QY 442 SerSerPhePheAlaAlaProAsnProGlyTrpAsnSerProHisGlyGlyIleGlyPro 461  
 DB 114 TCCAGTTTTTTTGGCCCCCAAAACCATATGGAATTCACCAACCTGGGGGATATATACCC 55  
 QY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTrpTyr 478  
 ||||||||||| :|||

DB 54 AACACTAGGCTCCGCCAGCCACCATATGCTTCTATCTGCGACTTACTAC 4  
 RESULT 10  
 ID AAK14426/C  
 ID AAK14426 standard; DNA; 567 BP.  
 XX  
 AC AAK14426;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe SEQ ID NO: 14417.  
 XX  
 KW Human: brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-0500667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 XX  
 PS Example 4: SEQ ID NO: 14417; 650pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is one of the probes of the  
 CC invention.  
 XX  
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1.78e-57 Length: 567  
 Score: 801.00 Matches: 151  
 Percent Similarity: 78.17% Conservative: 3  
 Best Local Similarity: 76.65% Mismatches: 3  
 Query Match: 30.95% Indels: 40  
 DB: 22 Gaps: 2  
 US-09-902-772-4 (1-478) x AAK14426 (1-567)  
 QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 321  
 DB 534 CTGACCTCATTTAGGAGAGTGGCCAGATCCAGCTTGGAGTCTCTCTGGAGCTCTGCG 475  
 QY 322 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGluPheLysMetThr 341  
 DB 474 GACAGCTCCACCTCAGCTGCATCACCCTGGAGGACCAACGAGGAGTCAAGTACG 415  
 QY 342 AspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTrp 361  
 ||||||||||| :|||

```

Db      414 GATCCGACGAGGTGGCCGGCGCTGGGGAGACGGGAAGCAAAACCAACATGACTAC 355
Qy      362 AsplysleuserAargAlaLeuAargTyrTyrAspLysAsnLleMetThrLysValHis 381
Db      354 GATAGCTCAGCGCGCCCTCGTACTACTATGCAAGACATCATGACCAAGTC--- 298
Qy      362 ProProgluserSerMetTyrLysTyrProSerAspLeuProTyrMetSerTyrHis 401
Db      297 -----CAT 295
Qy      402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
Db      294 GGGAAAGCGCTACGCTTACAGTTCAGCTTCACGGGATCGCCAGGCCCTCCAGCCAC 235
Qy      421 -----His 421
Db      234 CCCCCGAGTATCTCTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCCTATAC 175
Qy      422 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441
Db      174 GCCCACCACAGAAAGATGACTTGTGGCGCCACCTCCAGCCCTCCGCTGACATCT 115
Qy      442 SerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrGlyTyrLeuPro 461
Db      114 TCCAGTTTGTGCTGCCCCAAACCATGCTGGAATTCACCAACTGGGGGTATATACCC 55
Qy      462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
Db      54 AACACTAGGCTCCCGACCAAGCATATGCTTCTCATCTGGGCACTTACTAC 4
RESULT 11
AAK40160/c
ID      AAK40160 standard; DNA; 567 BP.
XX
AC      AAK40160;
XX
DF      06-NOV-2001 (first entry)
XX
DE      Human bone marrow expressed single exon probe SEQ ID NO: 14717.
XX
KW      Human; bone marrow expressed exon; gene expression analysis; probe;
KM      microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157276-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US00668.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488900/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
XX      analyzing gene expression in human bone marrow -
XX
PS      Example 4; SEQ ID NO: 14717; 658bp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
XX      probes which are derived from genomic sequences expressed in the human
XX      bone marrow. They can be used to measure gene expression in bone marrow

```

```

CC      samples, which may enable the improved diagnosis and treatment of cancers
CC      such as lymphoma, leukemia and myeloma. The present sequence is one of
CC      the probes of the invention.
XX
SQ      Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:
Pred. No.:      1.78e-57      Length:      567
Score:          801.00      Matches:      151
Percent Similarity: 78.17%      Conservative: 3
Best Local Similarity: 76.65%      Mismatches: 3
Query Match:    30.95%      Indels:      40
DB:             22      Gaps:      2

US-09-902-772-4 (1-478) x AAK40160 (1-567)
Qy      302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPheLeuLeuGlnLeuSer 321
Db      534 CTGACCTCATTAAGCAGTGGCCAGATCCAGCTTGGCGAGTCCCTCGACCTCGTGC 475
Qy      322 AspSerSerAsnSerAsnGlyIleThrTyrPheGlnGlyThrAsnGlyLysPheLysMetThr 341
Db      474 GACACTCTCACTCCAGCTGATCACCCTGGGAGGACCAACGGGAGTTCAAGATGACG 415
Qy      342 AspProAspGluValAlaArgArgTyrPheGlyLysSerLysProAsnMetAsnTyr 361
Db      414 GATCCCGACGAGGTGGCCGGCGCTGGGGAGAGCGGAGACCAAAACCAATGAACTAC 355
Qy      362 AsplysleuserAargAlaLeuAargTyrTyrTyrAspLysAsnLleMetThrLysValHis 381
Db      354 GATAGCTCAGCGCGCCCTCGTACTACTATGCAAGACATCATGACCAAGTC--- 298
Qy      382 ProProgluserSerMetTyrLysTyrProSerAspLeuProTyrMetSerTyrHis 401
Db      297 -----CAT 295
Qy      402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
Db      294 GGGAAAGCGCTACGCTTACAGTTCAGCTTCACGGGATCGCCAGGCCCTCCAGCCAC 235
Qy      421 -----His 421
Db      234 CCCCCGAGTATCTCTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCCTATAC 175
Qy      422 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441
Db      174 GCCCACCACAGAAAGATGACTTGTGGCGCCACCTCCAGCCCTCCGCTGACATCT 115
Qy      442 SerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrGlyTyrLeuPro 461
Db      114 TCCAGTTTGTGCTGCCCCAAACCATGCTGGAATTCACCAACTGGGGGTATATACCC 55
Qy      462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
Db      54 AACACTAGGCTCCCGACCAAGCATATGCTTCTCATCTGGGCACTTACTAC 4
RESULT 12
AAI20932/c
ID      AAI20932 standard; DNA; 567 BP.
XX
AC      AAI20932;
XX
DF      12-OCT-2001 (first entry)
XX
DE      Probe #10865 for gene expression analysis in human cervical cell sample.
XX
KW      Probe; human; microarray; gene expression; cervical epithelial cell;
KM      cervical cancer; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157278-A2.
XX

```

```
PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25; SEQ ID No 10865; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:
Pred. No.: 1.78e-57 Length: 567
Score: 801.00 Matches: 151
Percent Similarity: 78.17% Conservative: 3
Best Local Similarity: 76.65% Mismatches: 3
Query Match: 30.95% Indels: 40
DB: 22 Gaps: 2

US-09-902-772-4 (1-478) x AA120932 (1-567)
QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeuSer 321
DB 534 CTGACGTCATTAAGGCGAGTGGCCAGATCCAGCTTGGCAGTCCCTCGAGCTCTGTCG 475
QY 322 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGluPheLysMetThr 341
DB 474 GACGCTCCCACTCAGCTGCATCAGCTGGGAGGACCAACGAGGAGTTCAAGATGACG 415
QY 342 AspProAspGluValAlaArgArgTrpGlnArgLysSerLysProAsnMetAsnTyr 361
DB 414 GATCCCGAGGAGGGGCGGCGCTGGGAGGAGGAGGAGCAACCAACATGAATGACTAC 355
QY 362 AspLysLeuSerArgAlaLeuArgTrpTyrTyrAspLysAsnIleMetThrLysValHis 381
DB 354 GATNAGCTACAGCGCGGCGCTCGTACTACTACATGACAAGAACATGACCAAGTGC--- 298
QY 382 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401
DB 297 -----CAT 295
QY 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
DB 294 GGGAGAGCGCTACGCTTCAAGATTGACTTCCAGGAGTCCGCCAGCCCTCCAGCCAC 235
QY 421 -----His 421
DB 234 CCCCCGAGTATCTCTGTACAAAGTACCCCTCAGACCTCCGCTATGAGGCTCTATACAC 175
```

```
QY 422 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441
DB 174 GCCACCCACAGAAAGATGATTTGTGGCCGCCACCTCCAGCCCTCCGTCGACATCT 115
QY 442 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 461
DB 114 TCCAGTTTTTTTGTGCTGCCCAACCCATCTGGAATTTCACACAGTGGGGATATATACCC 55
QY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
DB 54 AACACTAGGCTCCGCCACGACCATATGCTTCTCATCTGCGGACACTTACTAC 4

RESULT 13
AA146176/c
ID AA146176 standard; DNA; 567 BP.
XX
XX AA146176;
AC
XX 17-OCT-2001 (first entry)
XX
XX Probe #14862 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray: human; placenta; antenatal diagnosis;
XX genetic disorder: ss.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID No 14862; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders.
XX
XX SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:
Pred. No.: 1.78e-57 Length: 567
Score: 801.00 Matches: 151
Percent Similarity: 78.17% Conservative: 3
Best Local Similarity: 76.65% Mismatches: 3
Query Match: 30.95% Indels: 40
DB: 22 Gaps: 2

US-09-902-772-4 (1-478) x AA146176 (1-567)
QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeuSer 321
DB 534 CTGACGTCATTAAGGCGAGTGGCCAGATCCAGCTTGGCAGTCCCTCGAGCTCTGTCG 475
```

Db 534 CTGACCTCATTAGGAGTGGCCAGATCCAGCTTTGGAGCTTCCTCCGAGCTCTGTCG 475  
 Qy 322 AsperSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPhelysMetThr 341  
 Db 474 GACAGCTCACTCAGCTGATCATCCCTGGAGAGCCAAAGGGAGTTCAAGATGACG 415  
 Qy 342 AsproAspGluValAlaIleArgTrpGlyGluArgLysSerLysProAsnMetAsnTyr 361  
 Db 414 GATCCCGACGAGTGGCCCGCGCTGGGGAGAGCGGAGCAACCAACCATGAACTAC 355  
 Qy 362 AspluSerArgAlaLeuArgTyrTyrAspLysAsnIleMetThrLysValHis 381  
 Db 354 GATAGCTCAGCGCGCTCGCTGCTACTACTATGCAAGAAATCATGACCAAGGTC--- 298  
 Qy 382 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401  
 Db 297 -----CAT 295  
 Qy 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420  
 Db 294 GGAAGCGCTACGGCTACAGATTCGATTCACAGGGATCCAGGCCCTCAGGCCAC 235  
 Qy 421 -----His 421  
 Db 234 CCCCCGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGCTCCTATAC 175  
 Qy 422 AlaHisProGlnLysMetAsnPheValAlaIleProHisProProAlaLeuProValThrSer 441  
 Db 174 GCCACCCACAGAAAGATGAATCTTGTGGCGCCACCCCTCAGGCCCTCCGCTGACATCT 115  
 Qy 442 SerSerPhePheAlaIleProAsnProTyrTyrAsnSerProThrGlyLysIleTyrPro 461  
 Db 114 TCCAGTTTTTTTGTGCCCCAACCCCACTGGAATTCACCACTGGGGGTATATACCCC 55  
 Qy 462 AsnThrArgLeuProAlaIleHisMetProSerHisLeuGlyThrTyrTyr 478  
 Db 54 AACACTAGGCTCCCGACCAACCATATGCTTCTCATCTGGGCACTTACTAC 4  
 RESULT 14  
 AA106643/c standard; DNA; 567 BP.  
 AC AA106643;  
 XX 09-OCT-2001 (first entry)  
 DE Probe #6634 used to measure gene expression in human breast sample.  
 XX  
 KW Probe: human; breast disease; breast cancer; development disorder; ss.  
 KW Inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 FN MO200157270-A2.  
 PD 09-AUG-2001.  
 XX  
 PF 29-JAN-2001; 2001WO-US00661.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-476286/51.  
 XX

PT Novel single exon nucleic acid probe used to measuring gene expression  
 PT in a human breast -  
 XX  
 PS Claim 25; SEQ ID NO 6634; 322pp; English.  
 XX  
 CC The present invention relates to novel single exon nucleic acid probes.  
 CC The present sequence is one such probe. The probes are useful for  
 CC measuring human gene expression in a human breast sample, where the probe  
 CC hybridises at high stringency to a nucleic acid expressed in the human  
 CC breast. The probes are useful for predicting, diagnosing, grading,  
 CC staging, monitoring and prognosing diseases of the human breast.  
 CC Particularly those diseases with polygenic aetiology. The diseases  
 CC include: breast cancer, disorders of development, inflammatory diseases  
 CC of the breast, fibrocystic changes, proliferative breast disease and  
 CC non-carcinoma tumours.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SO Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

## Alignment Scores:

Pred. No.:	1.78e-57	Length:	567
Score:	801.00	Matches:	151
Percent Similarity:	78.17%	Conservative:	3
Best Local Similarity:	76.65%	Mismatches:	40
Query Match:	30.95%	Indels:	2
DB:	22	Gaps:	2

US-09-902-772-4 (1-478) x AA106643 (1-567)

Qy 302. leuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 321  
 Db 534 CTGACCTCATTAGGAGTGGCCAGATCCAGCTTTGGAGCTTCCTCCGAGCTCTGTCG 475  
 Qy 322 AsperSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPhelysMetThr 341  
 Db 474 GACAGCTCACTCAGCTGATCATCCCTGGAGAGCCAAAGGGAGTTCAAGATGACG 415  
 Qy 342 AsproAspGluValAlaIleArgTrpGlyGluArgLysSerLysProAsnMetAsnTyr 361  
 Db 414 GATCCCGACGAGTGGCCCGCGCTGGGGAGAGCGGAGCAACCAACCATGAACTAC 355  
 Qy 362 AspluSerArgAlaLeuArgTyrTyrAspLysAsnIleMetThrLysValHis 381  
 Db 354 GATAGCTCAGCGCGCTCGCTGCTACTACTATGCAAGAAATCATGACCAAGGTC--- 298  
 Qy 382 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401  
 Db 297 -----CAT 295  
 Qy 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420  
 Db 294 GGAAGCGCTACGGCTACAGATTCGATTCACAGGGATCCCGCGCTCAGGCCAC 235  
 Qy 421 -----His 421  
 Db 234 CCCCCGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGCTCCTATAC 175  
 Qy 422 AlaHisProGlnLysMetAsnPheValAlaIleProHisProProAlaLeuProValThrSer 441  
 Db 174 GCCACCCACAGAAAGATGAATCTTGTGGCCCCAACCCCTCAGGCCCTCCGCTGACATCT 115  
 Qy 442 SerSerPhePheAlaIleProAsnProTyrTyrAsnSerProThrGlyLysIleTyrPro 461  
 Db 114 TCCAGTTTTTTTGTGCCCCAACCCCACTGGAATTCACCACTGGGGGTATATACCCC 55  
 Qy 462 AsnThrArgLeuProAlaIleHisMetProSerHisLeuGlyThrTyrTyr 478  
 Db 54 AACACTAGGCTCCCGACCAACCATATGCTTCTCATCTGGGCACTTACTAC 4  
 RESULT 15  
 ABS14215/c





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:18:59 ; Search time 64.3165 Seconds

(without alignments)  
2279.222 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588  
Sequence: 1 MASTRKALSVSHDSLF.....IYPTNRPAHMSHLCTTY 478

Scoring table:

BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n\_model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US0902772/runat\_23072003\_093657\_14904/app\_query.fasta.1.1294  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsun62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MOE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US0902772\_6CEN\_1.1\_55\_&runat\_23072003\_093657\_14904 -KCPU=6 -ICPU=3  
-NO\_MMAP -LARGEORDER -NBS\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCPUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2436	94.1	1528	4	US-08-878-177-3
2	2262.5	87.4	1447	4	US-08-878-177-1
3	1482	57.3	2938	2	US-08-343-443B-3
4	481	18.6	1752	4	US-09-360-779-1
5	442.5	17.1	1268	3	US-09-435-335-1
6	440.5	17.0	1604	1	US-08-344-579-1
7	440.5	17.0	1604	1	US-08-306-691B-43
8	440.5	17.0	1604	5	PCR-US93-06251-9
9	305.5	11.8	2667	2	US-08-469-412A-1
10	305.5	11.8	2667	4	US-09-021-715-1
11	301.5	11.6	328	2	US-08-343-443B-5
12	299.5	11.6	2064	3	US-08-875-944B-1

13	299.5	11.6	2064	4	US-09-116-049-3	Sequence 3, Appli
14	297.5	11.5	1905	4	US-09-055-113-2	Sequence 2, Appli
15	279.5	10.8	2410	2	US-08-780-835B-1	Sequence 1, Appli
16	279.5	10.8	2410	4	US-09-303-268-1	Sequence 1, Appli
17	279.5	10.8	2410	4	US-09-116-049-1	Sequence 1, Appli
18	266.5	10.3	2266	2	US-09-213-767-1	Sequence 1, Appli
19	257	9.9	5510	3	US-09-009-913-3	Sequence 3, Appli
20	245	9.5	5427	3	US-09-009-913-2	Sequence 2, Appli
21	245	9.5	5667	2	US-09-009-913-4	Sequence 4, Appli
22	230	8.9	2544	2	US-08-469-412A-6	Sequence 6, Appli
23	230	8.9	2544	4	US-09-021-715-6	Sequence 6, Appli
24	229	8.8	848	3	US-09-009-913-338	Sequence 338, App
25	213.5	8.2	852	4	US-09-020-956-44	Sequence 44, Appl
26	213.5	8.2	852	4	US-09-030-607-44	Sequence 44, Appl
27	213.5	8.2	852	4	US-09-605-785-44	Sequence 44, Appl
28	213.5	8.2	852	4	US-09-439-313-44	Sequence 44, Appl
29	213.5	8.2	852	4	US-09-352-616A-44	Sequence 44, Appl
30	213.5	8.2	852	4	US-09-232-149A-44	Sequence 44, Appl
31	213	8.2	2280	3	US-09-009-913-8	Sequence 8, Appli
32	213	8.2	2428	3	US-09-009-913-6	Sequence 6, Appli
33	213	8.2	2498	3	US-09-009-913-10	Sequence 10, Appl
34	211	8.2	1920	1	US-08-746-789A-1	Sequence 1, Appli
35	203	7.8	2975	1	US-08-368-281-1	Sequence 1, Appli
36	199	7.7	3240	1	US-08-368-281-3	Sequence 3, Appli
37	156	6.0	65042	4	US-09-784-316-3	Sequence 3, Appli
38	155.5	6.0	1364	1	US-08-306-691B-50	Sequence 50, Appl
39	155.5	6.0	1364	4	PCR-US93-06251-65	Sequence 65, Appl
40	129.5	5.0	50937	4	US-09-428-517-1	Sequence 1, Appli
41	124	4.8	1818	1	US-08-889-402-4	Sequence 4, Appli
42	124	4.8	1818	1	US-08-889-402-6	Sequence 6, Appli
43	122.5	4.7	2626	1	US-08-156-020-5	Sequence 5, Appli
44	121.5	4.7	2626	1	US-08-156-020-3	Sequence 3, Appli
45	121.5	4.7	2626	1	US-08-156-020-7	Sequence 7, Appli

#### ALIGNMENTS

RESULT 1  
US-08-878-177-3  
: Sequence 3, Application US/08878177  
: Patent No. 6294354  
: GENERAL INFORMATION:  
: APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et  
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
: TITLE OF INVENTION: the Proteins  
: FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001  
: CURRENT FILING DATE: 1997-06-18  
: NUMBER OF SEQ ID NOS: 7  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 3  
: LENGTH: 1528  
: TYPE: DNA  
: ORGANISM: c-ery gene, chicken DNA  
US-08-878-177-3

#### Alignment Scores:

Pred. No.: 3.69e-237  
Score: 2436.00  
Percent Similarity: 91.97%  
Best Local Similarity: 91.97%  
Query Match: 94.13%  
DB: 4  
Gaps: 2

US-09-902-772-4 (1-478) x US-08-878-177-3 (1-1528)

QY 1 MetlaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
DB 63 ATGCACAGCAGCTATTAAAGAGCATTTATCACTGGTGAAGACCAAGCTCTTTTGAG 122  
QY 21 CysAlaTyrgLysSerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
DB 123 TGTCCTACAGGATGCGCCCAAGCTTGCACAAAGACAGAAATGACAGCCTCTTCACAGTAA 182

QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 |||||  
 Db 183 TATGGGCAACATCAAAATGAGCCCGCGCTTCCCGCAGAGAGACTGGTATACAGAGCC 242  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 |||||  
 Db 243 CCGGCGAAGATTACATTAAAGATGAGTGTAAACCAACCAAGTTAAATGGGCAAGGAAAT 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 |||||  
 Db 303 TCACCTGATGACTGAGCGGTGGCAAAAGAGGAAATGGTTAGACAGTTGACACAAATTT 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrThr 120  
 |||||  
 Db 363 GGGATGAACATGTGAGAGCTACATGAGAAAGAGACATATCCGCCCTCCAAATATGACAAAC 422  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHisAlaArg 140  
 |||||  
 Db 423 AATGAACGAGAGATTATTGTGCCAGCAGATCTCTGATATGAGACACAGACCATGTACGG 482  
 QY 141 GlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 |||||  
 Db 483 CAGTGGCTGGAGTGGCGCAGTGAAGAGATGTGCTTCCAGACGTGGACATCTTGTGTTC 542  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 |||||  
 Db 543 CAGAACATTTGATGGAAAGAGTGTGTAAATGACCAAAAGATGACTTCCAGAGACTCAG 602  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGlnThrProLeu 200  
 |||||  
 Db 603 CCGAGCTATAACGAGATATCTCTCTGACACCTACCTACCTACCTACCTACCTACCTCTT 662  
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220  
 |||||  
 Db 663 CCACATTTGACTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722  
 QY 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240  
 |||||  
 Db 723 GCTGAAACACAGAGAGGAGCCACTTTATTTTCCAAATACATCATGTTTACCAGAGCA 782  
 QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTyr 260  
 |||||  
 Db 783 ACGCAAAAGATTAACAAAGAGCCAGATTTACCTTATGAGCAAGAGAGATGACAGCGTGG 842  
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280  
 |||||  
 Db 843 ACGAGTCACAGCCATCCACATCACTCAAAAGCTACCAACCATCATCTTCAACAGTGGCC 902  
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300  
 |||||  
 Db 903 AAAACAGAGACCCAGCGCTCTCACTAGATTCCTTATGAGATTTTGGACGACAGCAGC 962  
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuLeuGlnLeuLeu 320  
 |||||  
 Db 963 CGTCTTCCAATCCAGGAGAGGAGGAGATACAGATTAAGGCACTTCCACTGAGACTCTTG 1022  
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGluPheLysMet 340  
 |||||  
 Db 1023 TCGGACAGCTCCAACTCCAACTGATCATCTCGGAGAGGACCAAAATGGGAGTTAAAGATG 1082  
 QY 341 ThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsn 360  
 |||||  
 Db 1083 ACAGAACCTGATGAAGTGGCTCGGCGTGGGAGAGAGAGAAACCAAACTAACATGAC 1142  
 QY 361 TyrAspLysLysSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380  
 |||||  
 Db 1143 TATGACAACTCAGCCGCTGCACTTCTGCTACTACTATGACAAATAATATATGACTAAAGTT 1202  
 QY 380 ----- 380  
 Db 1203 CATGGTAAAGCTATGCTACAAATTTGATTTCCACGGAATCGCTCAGGCCCTCCAGCCT 1262

QY 381 HisProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSerTyr 400  
 |||||  
 Db 1263 CACCTCCAGATCATCTCATGTACAAATACCATCAGACCTCCCTACATGATGTTCTAC 1322  
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420  
 |||||  
 Db 1322 ----- 1322  
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440  
 |||||  
 Db 1323 CATGACACCCCGCAAGATGACACTTGTAGCTCCCATCCCGCTTTCGCCGTAC 1382  
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysTyr 460  
 |||||  
 Db 1383 TCATCCACCTTTTGTGCGCCCTAATCCATATGAGATTCACCAACGAGGACATCTAC 1442  
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478  
 |||||  
 Db 1443 CCCAATACCAAGGCTGCCAGCTGCTCATATGCTTCCCATCTTGGACCTACTAC 1496  
 RESULT 2  
 US-08-878-177-1  
 : Sequence 1, Application US/08878177  
 : Patent No. 6294354  
 : GENERAL INFORMATION:  
 : APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et  
 : TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
 : TITLE OF INVENTION: the proteins  
 : FILE REFERENCE: chugai selyaku kabushiki kaisha 5001  
 : CURRENT APPLICATION NUMBER: US/08/878,177  
 : CURRENT FILING DATE: 1997-06-18  
 : NUMBER OF SEQ ID NOS: 7  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 1  
 : LENGTH: 1447  
 : TYPE: DNA  
 : ORGANISM: C-11 gene, c-erg gene w/ deletion, chicken DNA  
 : US-08-878-177-1  
 Alignment Scores:  
 Pred. No.: 1, 28e-219 Length: 1447  
 Score: 2262.50 Matches: 430  
 Percent Similarity: 86.35% Conservative: 0  
 Best Local Similarity: 86.35% Mismatches: 1  
 Query Match: 87.42% Indels: 67  
 DB: 4 Gaps: 3  
 US-09-902-772-4 (1-478) x US-08-878-177-1 (1-1447)  
 QY 1 MetaLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerIleuPheGlu 20  
 |||||  
 Db 63 ATGCAAGCACTATTAAAGAAAGCTTATCAAGTGTGATGAAGACCAAGTCTTGTGAG 122  
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGlu 40  
 |||||  
 Db 123 TGTCCCTACAGATGCCGCCCACTTGCACAAAGACAAATAATGACAGCCTCTTCCAGTGAA 182  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 |||||  
 Db 183 TATGGGCAACATCAAAAGATGAGCCCGCGGCTTCCCGCAGCAGGAGCTGTTATACACAGCC 242  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 |||||  
 Db 243 CCGGCGAAGATTACATTAAAGATGAGTGTAAACCAACCAAGTTAAATGGGAGTTAAAGATG 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 |||||  
 Db 303 TCACCTGATGACTGAGCGGTGGCAAAAGAGGAAATGGTTAGACAGTTGACACAAATTT 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrThr 120  
 |||||  
 Db 363 GGGATGAACATGTGAGAGCTACATGAGAAAGAGACATATTCGCCCTCCAAATATGACAAAC 422

QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140  
 Db 423 AATGAAGAGAGATTATTGTCGCCAGCATCTCTGTTATGAGACAGACCATGTCAGG 482  
 QY 141 GluThrLeuGluThrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 Db 483 CAGTGGCTGGAGTGGAGTGAAGAGATGATGCTTCCAGACGTCGACATCTGTGTGTC 542  
 QY 161 GluAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGluArgLeuThr 180  
 Db 543 CAGACATGTATGGGAAAGAGTTGTGTAATGACCAAGATGACTTCCAGAGACTCAGC 602  
 QY 181 ProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200  
 Db 603 CCGAGCTATACGAGATATCTCTCTCTCACACCTACACTACCTCAGAGAGACA----- 656  
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGluIleAsnSerProArgLeuMetHis 220  
 Db 656 ----- 656  
 QY 221 AlaArgAsnThrGlyValAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240  
 Db 657 -----GGAGCCACTTATATTTTCCAAATACATCAGTTTACCCAGAGACA 701  
 QY 241 ThrGluArgIleThrThrArgProAspLeuProTyrGluGluAlaArgSerAlaTyr 260  
 Db 702 ACGCAAGAAATTAACAAGAGCCGATTTACCTTATAGCAGCAGAGAGATCGCTGG 761  
 QY 261 ThrSerHisSerHisProThrGluSerLysAlaThrGluProSerSerSerThrValPro 280  
 Db 762 ACGAGTCACAGCCATCCCTCAGTCAAGTCAAAAGCTACCAACATCATCTTCAACAGTCCCC 821  
 QY 281 LysThrGluAspGluArgProGluIleLeuAspProTyrGluIleLeuGluIleProThrSerSer 300  
 Db 822 AAACACAGAGACAGCCCTCTCAGTGTAGTCTTATCAGATCTTGGACCGACAGCAGC 881  
 QY 301 ArgLeuAlaAsnProGlySerGlyGluIleGluIleThrGluPheLeuLeuLeu 320  
 Db 882 CGTCTTGCAAATCCAGGAGATGGGACATACAGTATAGCGAGTTCCTACTGAGACTTCTG 941  
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTyrGluGluIleThrAsnGlyLeuPheLysMet 340  
 Db 942 TCGGACAGCTCCAACTCAATGTCATCCTGGAGGAGGACAAATGGGAGTTCAAGATG 1001  
 QY 341 ThrAspProAspGluValAlaArgArgTyrGlyLysLysSerLysProAsnMetAsn 360  
 Db 1002 ACAGACCCGTATGAAGTGGCTCGGCGTTGGGGAGAGAGAAACCAACCTTAACATGAC 1061  
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380  
 Db 1062 TATGACAAACTCAGCCGCTGCATCTGCTACTACTATGACAAATAATATATGACTAAAGTT 1121  
 QY 380 ----- 380  
 Db 1122 CATGGTAAACGCTATGCTACAAATTTGATTTCCACGGAATCGCTCAGCCCTCCAGCT 1181  
 QY 381 HisProProGluSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400  
 Db 1182 CACCTCTCAGAAATCATCATGTCATCAATATCCCATCAACCTCCCTCATGAGAGTTCTTAC 1241  
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGluAlaLeuGluPro 420  
 Db 1241 ----- 1241  
 QY 421 HisAlaHisProGluLysMetAsnPheValAlaProHisProAlaLeuProValThr 440  
 Db 1242 CATGCAACACCCAGAGATGAACTTGTAGTCCCATCCCTCGCTTGGCCGTAACC 1301  
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrArgLysGlyIleTyr 460  
 Db 1302 TCATCCACGCTTTTGTGCTGCCCTAAATCCATACTGGAATTCACCAACTGGAGGCACTTAC 1361  
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478

Db 1362 CCCAATACAGGCTGCCAGCTGCATATATGCTTCCATCTGTGACACTACTAC 1415  
 RESULT 3  
 US-08-343-443B-3  
 : Sequence 3, Application US/0834343B  
 : Patent No. 5968734  
 GENERAL INFORMATION:  
 : APPLICANT: Aurias, Alain  
 : APPLICANT: Delatize, Olivier  
 : APPLICANT: Desmaze, Chantal  
 : APPLICANT: Melot, Thomas  
 : APPLICANT: Peter, Martine  
 : APPLICANT: Plooungastel, Beatrice  
 : APPLICANT: Thomas, Gilles  
 : TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF  
 : TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL  
 : TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS  
 : TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID  
 : NUMBER OF SEQUENCES: 129  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Weiser & Associates  
 : STREET: 230 South Fifteenth Street  
 : CITY: Philadelphia  
 : STATE: PA  
 : COUNTRY: USA  
 : ZIP: 19102  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: AEDIT 1.0 DOS text editor  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/343,443B  
 : FILING DATE: 18-NOV-1994  
 : CLASSIFICATION: 514  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/FR93/00494  
 : FILING DATE: 19-MAY-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: FR 92/06123  
 : FILING DATE: 20-MAY-1992  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Weiser, Gerard J.  
 : REGISTRATION NUMBER: 19,763  
 : REFERENCE/DOCKET NUMBER: 989,6121P  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 215-875-8383  
 : TELEFAX: 215-875-8394  
 : INFORMATION FOR SEQ ID NO: 3:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2938 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 143..1498  
 : US-08-343-443B-3  
 Alignment Scores:  
 Pred. No.: 4.88e-140 Length: 2938  
 Score: 1482.00 Matches: 291  
 Percent Similarity: 68.65% Conservative: 55  
 Best Local Similarity: 57.74% Mismatches: 80  
 Query Match: 57.26% Indels: 78  
 DB: 2 Gaps: 9  
 US-09-902-772-4 (1-478) x US-08-343-443B-3 (1-2938)  
 QY 1 MetalaserThrIleLysGluAlaLeuSerValValSerGluAspGluInsLeuPheGlu 20

```

143 ATGACGGGACTATTAAAGAGGCTCTGCGTGGAGCAGACCACTCCCTTTGAC 202
21 CysAlaTyrGly----SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
203 TCACAGTACGAGGAGCGACGCTCCCAAGCGCCAGCATGCTCCGGGAGCTCT 262
40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGln 59
263 GACATACGGGAGCGCCCAACATACACCCCTCCACACACAGCATGATCATCAG 322
60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
323 ---CCAGTAGGAGTCAACGCTCAACGCGGAGTAT-----GACCAATGAATGATCCAG 373
80 AsnSerProAspSerCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
374 GAGTCTCCGCTGCTGAGCTGAGCGTTAGCAATGCAAGAGTGGGCGGAGCGAGTCC 433
100 ValGlyMetAsnTyrGlySerTyrMetGluGlnLysHis---IleProProAsnMet 118
434 AACCCCATGACATACACAGCATATATGACAGAAAGATGGCCCCCTCCCTCCACATG 493
119 ThrThrAsnGlnArgValIleValProAlaAspProThrLeuTyrSerThrAspHis 138
494 ACCACCAACGAGAGAGATCATGCTCCCGCAGACCCACACATGTCAGACAGAGCAT 553
139 ValArgGlnTyrPleuGlnTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeu 158
554 GTGAGGCAATGCTGGAGTGGTGAATAAGATGATGCTTATGATGATGACATGCAATCC 613
159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
614 TTTTTCGAGACATGATGAGCAAGAACTGTGTAATAAGAACAGAGAGATCTCTCCG 673
179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisIleHisTyrLeuArgGluThr 198
674 GCCACCAACCTCTACACACAGCAAGTGTGTGCACACCTCACTTACCTCAGGGAAGT 733
199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
734 TCACATG----- 739
219 MethIleAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
740 -----CTGGCCCTATATATCAACTCCACAC 766
239 GluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSer 258
767 GACCAATCTCTACGATGAGTGCATAAAGAACACCTTCTTATGACTCAGTCAGAAAGAGA 826
259 AlaTyrThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
827 GCATGGGGCAATATACATGATTCGGCTCAACAAAAGTCTCCCTTGGAGGGGCACAA 886
278 ThrValProLysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297
887 ACGATCAGTAAAGATACAGACAGACAGCCCGACAGATCGGTACAGATCTGGGGCCG 946
298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeu 317
947 ACCGACAGTCGCTCAGCAACCCCGAAGCGGGAGATCCAGCTGTGCAATCTCTCTG 1006
318 GluLeuLeuSerAspSerSerAsnSerAsnCysIleThrThrPheGluGlyThrAsnGlyGlu 337
1007 GAGCTGCTCTCCGACAGCGCCACAGCGCATGTATCACTGGGAGGGAGCCCAAGCGGAG 1066
338 PheLysMetThrAspProAspGluValAlaArgArgTyrPheGlyLysArgSerLysPro 357
1067 TTCAAAATGACGAGCGCCGATGAGTGGCCAGCGGCGTGGGGCGAGGAAAGCAACGCC 1126
358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377

```

```

Db 1127 AACATGATTAACAGAACTAGCGCGGCTCCGTTATTAATGATGATAAAACATTAATG 1186
Oy 378 ThrLysValHisProProGlnLysSerSerMetTyrLysTyrProSerAspLeuProTyrMet 397
Db 1187 ACCAAAGT----- 1195
Oy 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417
Db 1196 -----CACGGCAAAAGATATGCTTACAAATTTGACTTCCACGGCATTTGCCAGGCT 1246
Oy 418 LeuGlnPro----- 420
Db 1247 CTGACGACCAATCCGACCGAGCTGTCATGTAACAGTACCCCTTGACATCTCTATCATG 1306
Oy 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeu 437
Db 1307 CCTTCTACCATGCCCCACACAGCAAGGTGAACTTTGCTCCCTCCCATCCATCTCCATG 1366
Oy 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGly 457
Db 1367 CCTGTCACTTCTCCAGCTTCTTGGAGCCGATACACATACATGAGCTCCGCCACGAGG 1426
Oy 458 GlyIleTyrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474
Db 1427 GGAATTCACCCCAACCCCAACGTCGCCGATCTCAACACCAAGTGCCTTCACTTA 1486
Oy 475 GlyThrTyrTyr 478
Db 1487 GCGAGCTACTATG 1498

RESULT 4
US-09-360-779-1
: Sequence 1, Application US/09360779
: Patent No. 6268216
: GENERAL INFORMATION:
: APPLICANT: Demeris, Evan S.
: APPLICANT: Eyodoio, Dmitry V.
: APPLICANT: Hendricks, Timothy J.
: TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
: TITLE OF INVENTION: Useful In the Treatment of Neurological Diseases
: FILE REFERENCE: CASE-03828
: CURRENT APPLICATION NUMBER: US/09/360, 779
: CURRENT FILING DATE: 1999-07-26
: EARLIER APPLICATION NUMBER: 60/094, 264
: EARLIER FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: patentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1752
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (112)..(1131)
US-09-360-779-1

Alignment Scores:
Pred. No.: 5,54e-39 Length: 1752
Score: 481.00 Matches: 122
Percent Similarity: 44.73% Conservative: 18
Best Local Similarity: 38.98% Mismatches: 68
Query Match: 18.59% Indels: 105
DB: 4 Gaps: 10

US-09-902-772-4 (1-478) x US-09-360-779-1 (1-1752)
Oy 224 ThrGlyAlaLathrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArg 243
Db 333 ACGGGCGGATGCGCA-----CAAGCGCAAGCTGACCTGCCCGCGTGGGGTCCCCC 392
Oy 244 IleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaThrThrSerHis 263
Db 393 ATCTCCACCGGCCCA-----GTCCCGGCG 416

```

```

QY 264 SerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGlu 283
DB 417 AGCGATGAGACAGACGGCCACCTCCAGCCCTGCTGATCAACTGTAACCTACCAATCC 476
QY 284 AspGlnArg-----ProGlnLeu-AspProTyrglnIleLe 295
DB 477 CGTGGAGATGGTCTTTTAAAGAAAGAGAGCCGAGCGGGGGCG-----CT 527
QY 295 uGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPh 315
DB 528 GAGCGCTGGGTACAGAA-----GGCAGCGGGCAGATCCAGTTGTGGCAGTT 575
QY 315 eLeuLeuGlnLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGlnIleTrpAs 335
DB 576 TCTACTGAGAGCTGCTGGCAGACCGCGGAACCGCGCTGCATCCCGTGGAGAGCGGCCA 635
QY 335 nGlyGluPheLysMetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSe 355
DB 636 CGCGGAGTTCAAGCTCAACCGCCGACGAGGTGGCGCGAGCTGGGGCGAGCCAAAG 695
QY 355 rLyAspProAsnMetAsnTyAspLysLeuSerArgAlaLeuArgTyTrpAspLysAs 375
DB 696 CAAGCCCAATATGAACACTAGCAGCAAGCTTAAGTGAAGCTGGCTACTACTACGCAAAA 755
QY 375 nIleMetThrLysValHisProProGlnSerSerMetLysTyTrpProSerAspLeuPr 395
DB 756 CATCATGACGAAAGTG-----HisAlaHisPro----- 771
QY 395 oTyMetSerSerTyHisGlyLysArgTyAlaTyLysPheAspPheHisGlyIleAl 415
DB 772 -----CAGGCAAGCGCTACGCTTACCGCTTGTGACTTCCAGGGCCGCGC 815
QY 415 aglnAlaLeuGlnPro-----HisAlaHisPro----- 424
DB 816 ACAAGCTTGGCAGACACCCCGCGACGCGCCAGCGCGCTGCGCGCGCAGAGCGC 875
QY 424 ----- 424
DB 876 AGCGCGCGCCAGATGGCGCACTTTACAAAGCTCCCGGCTGGTGGCTCCATGCCCTT 935
QY 425 -----GlnLysMetAsnPheValAlaAlaProHisProProAlaLeuProValAla 441
DB 936 CCGCGGCTCTCCAACTCAACTTATGGCAGCTCGCGCGCTGGCGCGCTGGCTG 995
QY 441 rSerSerPhePheAlaAlaProAsnProTyTrpAsnSerPro----- 455
DB 996 CTCT-----TACTGGCTGGTCCCAACGCCAGCGCGCTGC 1031
QY 456 -----ThrGlyGlyIleTyTrpAsnThrArgLeuPro----- 466
DB 1032 CGCGCGCAACCGCTGCTACCAACCCCGGGCTTGCAGCCCTCCCGGGCCCTTGG 1091
QY 467 -AlaAlaHisMetProSerHisLeuGlyIleTyTrp 478
DB 1092 CGCGGTGGCGCGCTTCGCACCTTGGGGGCTATTAT 1128

```

RESULT 5  
US-09-435-335-1  
Sequence 1, Application US/09435335  
Patent No. 6384204  
GENERAL INFORMATION:  
APPLICANT: Demeiris, Evan S.  
APPLICANT: Pyodoro, Dmitry V.  
APPLICANT: Hendricks, Timothy J.  
TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds  
FILE REFERENCE: CASE-04027  
CURRENT APPLICATION NUMBER: US/09/435,335  
EARLIER FILING DATE: 1999-11-05  
EARLIER APPLICATION NUMBER: 09/360,779  
NUMBER OF SEQ ID NOS: 23

```

: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1752
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (112)..(1131)
US-09-435-335-1

Alignment Scores:
Pred. No.: 5,54e-39 Length: 1752
Score: 481.00 Matches: 122
Percent Similarity: 44.73% Conservative: 18
Best Local Similarity: 38.98% Mismatches: 68
Query Match: 18.59% Indels: 105
DB: 4 Gaps: 10

US-09-902-772-4 (1-478) x US-09-435-335-1 (1-1752)
QY 224 ThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyProGluAlaThrGlnArg 243
DB 333 AGGGGGGGTATCGGACACCCCAAGCCAAAGCTACGTCGCCCGCGGTCCCGCC 392
QY 244 IleThrThrArgProAspLeuProTyTrpGlnAlaAlaArgSerAlaTrpThrSerHis 263
DB 393 ATCTCCACCGCCCA-----GTCCCGCGC 416
QY 264 SerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGlu 283
DB 417 AGCGATGAGACAGACGGCCACCTCCAGCCCTGCTGATCAACTGTAACCTACCAATCC 476
QY 284 AspGlnArg-----ProGlnLeu-AspProTyrglnIleLe 295
DB 477 CGTGGAGATGGTCTTTTAAAGAAAGGAAAGCCGAGCTGGGGCGC-----CT 527
QY 295 uGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPh 315
DB 528 GAGCGCTGGGTACAGAA-----GGCAGCGGGCAGATCCAGTTGTGGCAGTT 575
QY 315 eLeuLeuGlnLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGlnIleTrpAs 335
DB 576 TCTACTGAGAGCTGCTGGCAGACCGCGGAACCGCGCTGCATCCCGTGGAGAGCGGCCA 635
QY 335 nGlyGluPheLysMetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSe 355
DB 636 CGCGGAGTTCAAGCTCAACCGCCGACGAGGTGGCGAGCTGGGGCGAGCCAAAG 695
QY 355 rLyAspProAsnMetAsnTyAspLysLeuSerArgAlaLeuArgTyTrpAspLysAs 375
DB 696 CAAGCCCAATATGAACACTAGCAGCAAGCTTAAGTGAAGCTGGCTACTACTACGCAAAA 755
QY 375 nIleMetThrLysValHisProProGlnSerSerMetLysTyTrpProSerAspLeuPr 395
DB 756 CATCATGACGAAAGTG-----HisAlaHisPro----- 771
QY 395 oTyMetSerSerTyHisGlyLysArgTyAlaTyLysPheAspPheHisGlyIleAl 415
DB 772 -----CAGGCAAGCGCTTACGCTTACCGCTTGTGACTTCCAGGGCCGCGC 815
QY 415 aglnAlaLeuGlnPro-----HisAlaHisPro----- 424
DB 816 ACAAGCTTGGCAGACACCCCGCGACGCGCCAGCGCGCTGCGCGCAGAGCGC 875
QY 424 ----- 424
DB 876 AGCGCGCGCCAGATGGCGCACTTTACAAAGCTCCCGGCTGGTGGCTCCATGCCCTT 935
QY 425 -----GlnLysMetAsnPheValAlaAlaProHisProProAlaLeuProValAla 441
DB 936 CCGCGGCTCTCCAACTCAACTTATGGCAGCTCGCGCGCTGGCGCGCTGGCTG 995
QY 441 rSerSerPhePheAlaAlaProAsnProTyTrpAsnSerPro----- 455

```

```

Db      996  CTCT-----TACTGGCTGTCCCAAGCCACCGCGCTGC 1031
Oy      456  -----ThrlgyllyltyrProAsnThrArgLeuPro----- 466
Db      1032 CGCGCCGCGCTGCGCTGTACCCCAAGCCGCGCTGTGAGCCCTCCGCGGCTTTGG 1091
Oy      467  -AlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db      1092 CGCGGTGGCGCGCGCTTGCGCACTTGGGGGCTCATTTAT 1128

RESULT 6
US-09-344-579-1
; Sequence 1, Application US/09344579
; Patent No. 6054316
; GENERAL INFORMATION:
; APPLICANT: Brenda F. Baker
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF ETS-2 EXPRESSION
; FILE REFERENCE: RTS-0063
; CURRENT APPLICATION NUMBER: US/09/344,579
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (291)..(1700)
US-09-344-579-1

```

```

Alignment Scores:
Pred. No.: 6,71e-35      Length: 2268
Score: 442.50           Matches: 133
Percent Similarity: 39.50%      Conserves: 40
Best Local Similarity: 30.37%    Mismatches: 108
Query Match: 17.10%           Indels: 158
                                Gaps: 16
US-09-902-772-4 (1-478) x US-09-344-579-1 (1-2268)

```

```

Oy      122  GUAARGATGVALIIEVALProAlaSPProThrLeuThrSerThrAspHisValArgGln 141
Db      570  CAGGGGCGCGCTGGGATTCGCAAGAACCCCTGGCTGTGAGTGAACGTGACAGAGTTC-- 629
Oy      142  TTPLeuGlnThrPalValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db      630  TGCGTCTCTGGGCGCCACCATGATGATCTGTGTGAACGTGAATGTGACAGAGTTC-- 686
Oy      162  AsnIleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db      687  GGCATGATGAGCGCATGCTGTGTAACTTGGCAAGAACGCTTTCGAGACTGGCACT 746
Oy      182  SerTyrAsnAlaAspIleLeuLeuSerHisLeu----- 192
Db      747  GACTTGTGGTGACATCTCTGTGGAACATCTGGCAAAATGACAAAGAAACCAAGAA 806
Oy      193  -----HistyrLeuArg 196
Db      807  AAGACAGACATCAATATGAAGAAATTCACACCTCACCCTCCGTTCTCTCATTTGATTAA 866
Oy      197  GluThr----- 198
Db      867  AGCAATACATTAGTTTGGACAGACAGCGCCCTATGGAATGACAGACACAGAAATTAC 926
Oy      199  -----ProLeu-ProHisLeuThrSerAspAspVa 208
Db      927  CCCAAAGCGGCTCTCTGAGACAGCATGTGCGGCTCCACACCCACGCTACTGAGCTCT 986
Oy      208  LAspLysAlaLeuGlnAsnSerProArgLeu----- 218
Db      987  GAGCAGAGATTTCAGATGTTCCCAAGTTCGGCTCAGCTCCGTCACGCTACCTACTGC 1046

```

```

Oy      219  -MetHisAlaArgAsnThrGlyAlaThrPheIle-----Ph 231
Db      1047  TCTGTACATCAGGAGCTTCCACGACGACCACTTGATTTGCTCACCACAAATTCTGGAGCT 1106
Oy      231  eProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgProAspLeuPr 251
Db      1107  CCCAAAGACACGACT---CCCTTGAGAACGGTGGGACAGCTTCAGAGAGCTCAGACTCC 1163
Oy      251  oTyrGluGln-----AlaArgArgSerAlaTrp----- 260
Db      1164  CTCCTCCAGTCTGGAAACAGCAGTCTCTGTGCTGATGACAGCGGTCTCTCTCCTTC 1223
Oy      261  -----ThSerHisSerHisProThrGlnSe 269
Db      1224  GAGAGCTTCGAAGATGACGTGACGACGCTCTCTGCTCATTAAGCCACCATGCTCTTTC 1283
Oy      269  TlysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgProGlnLe 289
Db      1284  AAGATTACA-----TCCAAGAGAGGAGTGACCCAGTG--CAGCAAGGCAAAACAGTTAT 1336
Oy      289  uAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGly 309
Db      1337  ACTGCAAGCTGTGTGGCGGCTTCACA-----GGAAGTGACC 1375
Oy      309  nIleGlnLeuThrPheLeuLeuGluLeuSerAspSerSerAsnSerAsnGly 329
Db      1376  TATTCAGCTGTGGAGATTTCCTCGAGCTCTCATACAGAAATCCGCGCATTCATCAT 1435
Oy      329  eThrTTPGluGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAlaArg 349
Db      1436  CAGCTGAGACTGGAACAGGATGTGAAGCTTAAGCTCGCGGACCCCATAGGTGGCCCGC 1495
Oy      349  gTTPGluGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArg 369
Db      1496  GTGGGGAAGAGGAAATTAAGCCCAAGATGATACGAGAGCTGAGCGGCGCTTACG 1555
Oy      369  gTyrTyrTyrAspLysAsnIleMetThrLysValHisProGluSerMetTyrLys 389
Db      1556  CTACTATTACAGACAAAGAACATCATCATCAAGAGAGC----- 1589
Oy      389  sTyrProSerAspLeuProTyrMetSerSerTyrHisGlyLysArgTyrLysPhe 409
Db      1590  -----TCGGGGAAGCGCTACGTGACCCCTT 1615
Oy      409  e-----AspPheHisGlyIleAlaGln 416
Db      1616  CGTGTGGAGCTCCAGAACTTGCTGGGCTTCACGCCCGGAGAGACTGCACGCTCTG-- 1673
Oy      416  nAlaLeuGlnProHisAla-----HisProGlnLysMetAsnPh 429
Db      1674  -GGGCTCAGCGCCGACACAGGAGAGCTAGGTCCCGGAGACACCTGAGCGG----- 1724
Oy      429  eValAlaPro-----HisProAlaLeuPro 438
Db      1725  ---GCCCAAGGCTGTGAGCTGAGTGAGGAGGCCATCTGACCACTGCT 1772

RESULT 7
US-08-306-691B-43
; Sequence 43, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavoragna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.

```





```

; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1604 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-06251-9

Alignment Scores:
Pred. No.: 6.09e-35      Length: 1604
Score: 440.50           Matches: 115
Percent Similarity: 44.21%      Conservative: 53
Best Local Similarity: 30.26%   Mismatches: 102
Query Match: 17.02%           Indels: 111
DB: 5                      Gaps: 8

US-09-902-772-4 (1-478) x PCT-US93-06251-9 (1-1604)
QY 119 ThrThrsnngluArGArGValIlleAlProAlaAspProThrLeuTrpSerThrAspHis 138
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 447 ACTAAGAACAGCAACGACGCTGGGATCCCAAAAGACCCCGGAGTGACAGAAACCAT 506
QY 139 ValArgGlnTrpLeuGluTrpAlaValIleYslGluTrpLeuProAspValAspIleLeu 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 507 GTTGCGAGCTGGGATGGCTGGCTGATGATTCAGACCCGCAAGAGTGAGCTTCAG 566
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuGlyMetThrLysAspAspPheGlnArg 178
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 567 AAGTTC---TGTATGATGGAGGACGCCCTCTGCCCTGGGTAAAGACTGCTTCTCGAG 623
QY 179 LeuThrProSerTyrrAsnAlaAspIleLeuLeuSerHisLeuHisTyrrLeuArgGluThr 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 624 CTGGCCCCAGACTTGTGGGGACATCTTATGGAAACATCTAGAGATCCCTCAGAAAGAG 683
QY 199 ProLeuPro-HisLeuThrSerAspAspValAspLysAlaLeuGlnAsn----- 214
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 684 GATGTGAACCATATVCAAGTTAATGAGTCAACCCACCTATCCAGATCCCGCTATACC 743
QY 215 -----SerProArgLeuMetHisAlaArgAsnTh 224
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 744 TCGGATTACTTATTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 803
QY 224 rGlyGlyAlaThrPheIle----- 230
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 804 TCAGAGCCC-AGCTTCATTCACAGAGCTCTATCAGACGCTCCATCCATCCAGCTCGAAGA 862
QY 231 -----PheProAsnThrSerVal----- 236
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 863 GCTCCTCTCCCTCAAGATGAGAAATGACTACCTACCTCGGTATCTCCGAGACCTCTCCA 922
QY 237 -----TyrProGluAlaThrGlnArgGlyLeuThrTrpArgProAs 249
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 923 GACAGACACCTTGACGAATGACTACTTTCATCAAAACAGAAAGTCTCACCCACAGACAA 982
QY 249 PLeuProTyrrGluGlnAlaArgArg----- 257
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 983 CATGTGCATGGGAGGACGACGCTGTGTAACCTGGGGGCGCAGACTCTTTGAAGCAT 1042
QY 258 -SerAlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSe 277
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1043 AGAAGAGCTAGATAGTGTGATCGCCCTACCCAGATCTCTGGAGCGCAGATCTTTCAA 1102
QY 277 rThrValProLys-----ThrGluAsp----- 284
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1103 CAGCTCTCAGAGGTGTTCTCTCTATAGACAGCTTCGACTCAGAGGACTATCCGGCTGCCCT 1162
QY 285 -----GlnArgProGlnLeuAs 290
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1163 GCCCAACCAAGGCCCAAGGACACTTCAAGAGACTATGTGGGAGCGCTGCTGACCTCAA 1222
QY 290 pProTyrrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnI 310

```

```

Db 1223 TAAGACAAAGCCTCTCATCTCTGCTGCTGCTAGCTGGCTACACAGGACGACGTGACCAT 1282
QY 310 eGlnLeuTrpGlnPheLeuGluLeuLeuSerAspSerSerSerAsnGlyIleTh 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1283 CCAGCTATGGCAGTTCTTCTCGAATTAATCTAGTAATATCCGTGACTGTTTATTCAG 1342
QY 330 rTrpGluGlyThrAsnGluLuphLeuMetThrAspProAspGluValAlaArgArgTr 350
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1343 CTGGACAGAGATAGCTGGGAGATTCAAACTTCTGACCCAGATAGAGTGGCGCAGAGATG 1402
QY 350 pGlyGluArgLysSerLysProAsnMetAsnTyrrAspLysLeuSerArgAlaLeuArgTy 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1403 GGGAAAGAGGAAAAAACAACCTTAGATGATATATAGAGAACTGAGCCGTGCGCTACCTA 1462
QY 370 rTyrrAspLysAsnIleMetThrLysValHisProProGluSerSerMetTyrrLysTy 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1463 CTATTACGACAAATATCATCCACAAAGACA----- 1493
QY 390 rProSerAspLeuProTyrrMetSerTyrrHisGlyLysArgTyrrAlaTyrrLysPhe 409
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1494 -----GGCGGGAACGCTAGCTGATACCGCTTT 1520

```

```

RESULT 9
US-08-469-412A-1
; Sequence 1, Application US/08469412A
; Patent No. 5856125
;
GENERAL INFORMATION:
; APPLICANT: MayoThalassitis, George J.
; APPLICANT: Blair, Donald G.
; APPLICANT: Fisher, Robert J.
; APPLICANT: Beal Jr., Gregory J.
; APPLICANT: Athanaslou, Meropi A.
; APPLICANT: Sgouras, Dionysios N.
; TITLE OF INVENTION: The ERF Genetic locus and its products
; NUMBER OF SEQUENCES: 16
;
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,412A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,350
; REFERENCE/DOCKET NUMBER: 015280-229000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2667 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 123..1769
; OTHER INFORMATION: /note="human ERF (ETS2 Repressor
; OTHER INFORMATION: Factor) cDNA"
US-08-469-412A-1

```

## Alignment Scores:

Pred. No.: 6,64e-21 Length: 2667  
 Score: 305.50 Matches: 88  
 Percent Similarity: 50.91% Conservative: 24  
 Best Local Similarity: 40.00% Mismatches: 64  
 Query Match: 11.80% Indels: 47  
 DB: 2 Gaps: 7

US-09-902-772-4 (1-478) x US-08-469-412A-1 (1-2667)

```

QY 266 ProThrGlnSerLysAlaThrGlnProSerSerSerThraValProLysThrGluAsp--G 285
DB 74 CCGGAGAGCCCGCGCCGGAATCGGGGGCGCTTCGCCCGCGCCCGCCAGCATGAAGACCCC 133
QY 285 lnaArgProGlnLeuAspProTyrglnlleuGlnProThrsSerArgLeuAlaAsp 305
DB 134 GCGGAGACACAGGAGTTTGGCTTCCCGGATT--GGGCGCTACA--AGCCAGAG-TCGTCCC 186
QY 305 roGlySerGlyGlnlleGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSera 325
DB 187 CTGGCTCAAGCAGCATTCAGCTGCGACCTTATCTCGAGCTCTCGGGAAGAGAGAGT 246
QY 325 snSerAsnCysIleThrTrpGlnGlyThraSnglyGluPheLysMetThrAspProAspG 345
DB 247 ACCAGGCGCTATTCCTGGCAGGAGGAGCTACGAGGGAATTGCTCATCAAGACCTGATG 306
QY 345 luValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyraSplyLeus 365
DB 307 AGGTGGCCCGGCTGTGGGGCGCTTGCAGATGCAAGCCCGCATGATTAACGACAAGCTGA 366
QY 365 eArGAlaLeuArgTyrrTyrrAspLysAsnIleMetThrLysValHisProProGlns 385
DB 367 GCGGCGCCCTCGCTATTACTATTAACAAGCCGATTCGACAAAGC----- 413
QY 385 erSerMetTyrrLysTyrrProSerAspLeuProTyrrMetSerSerTyrrHisGlyLysArgT 405
DB 414 -----AAGGGAAACGGT 426
QY 405 yAlaTyrrLysPheAspPheHisGlylleAlaGlnAlaLeuGlnProHis----- 421
DB 427 TCACCTACAAAGTTCAATTTCAACAACAGCTGTGTCGCAATTACCCATTGATGATG 486
QY 422 -----AlaHisProGlnLysMetAsnPheValAlaProHisProProAla 437
DB 487 GGTGGCTGGGGGTGCGACGTGCCACAGT-----GCCCGCCAG 525
QY 437 euProValThrSerSerSerPhePheAlaAlaProAsnProTyrrTrpAsnSerProThrg 457
DB 526 TGCCGTGGGTGTAGCCACTTCCGCTTCCCTCC-----TCACGCGCCCTCCG 573
QY 457 lylGlylleTyrrProAsnThrArgLeuProAlaAlaHisMetProSer---HisLeu 474
DB 574 A-GGTGCTGTCCCCACCG-AGGACCCCGCTCACACACGCGTCTTCATCTT 627
  
```

RESULT 10  
 US-09-021-715-1  
 : Sequence 1, Application US/09021715  
 : Patent No. 6194547

## GENERAL INFORMATION:

APPLICANT: Mavrothalassitis, George J.  
 Inventor: Blair, Donald G.  
 Fisher, Robert J.  
 Beal Jr., Gregory J.  
 Athanasios, Meropi A.  
 Sgouras, Dionysios N.  
 TITLE OF INVENTION: The ERF Genetic Locus and Its Products  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Two Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA

```

? ZIP: 94111-3834
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/021,715
? FILING DATE: 10-Feb-1998
? CLASSIFICATION: <unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Garrett-Mackowski, Eugenia
? REGISTRATION NUMBER: 37,330
? REFERENCE/DOCKET NUMBER: 015280-229000
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 576-0200
? TELEFAX: (415) 576-0300
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2667 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 123..1769
? OTHER INFORMATION: /note= "human ERF (ERF2 Repressor
? Factor) cDNA"
? SEQUENCE DESCRIPTION: SEQ ID NO: 1:
?
? US-09-021-715-1
?
? Alignment Scores:
? Pred. No.: 6,64e-21 Length: 2667
? Score: 305.50 Matches: 88
? Percent Similarity: 50.91% Conservative: 24
? Best Local Similarity: 40.00% Mismatches: 64
? Query Match: 11.80% Indels: 47
? DB: 4 Gaps: 7
?
? US-09-902-772-4 (1-478) x US-09-021-715-1 (1-2667)
?
QY 266 ProThrGlnSerLysAlaThrGlnProSerSerSerThraValProLysThrGluAsp--G 285
DB 74 CCGGAGAGCCCGCGCCGGAATCGGGGGCGCTTCGCCCGCGCCCGCCAGCATGAAGACCCC 133
QY 285 lnaArgProGlnLeuAspProTyrglnlleuGlnProThrsSerArgLeuAlaAsp 305
DB 134 GCGGAGACACAGGAGTTTGGCTTCCCGGATT--GGGCGTACA--AGCCAGAG-TCGTCCC 186
QY 305 roGlySerGlyGlnlleGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSera 325
DB 187 CTGGCTCAAGCAGCATTCAGCTGCGACCTTATCTCGAGCTCTCGGGAAGAGAGAGT 246
QY 325 snSerAsnCysIleThrTrpGlnGlyThraSnglyGluPheLysMetThrAspProAspG 345
DB 247 ACCAGGCGCTATTCCTGGCAGGAGGAGCTACGAGGGAATTGCTCATCAAGACCTGATG 306
QY 345 luValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyraSplyLeus 365
DB 307 AGGTGGCCCGGCTGTGGGGCGCTTGCAGATGCAAGCCCGCATGATTAACGACAAGCTGA 366
QY 365 eArGAlaLeuArgTyrrTyrrAspLysAsnIleMetThrLysValHisProProGlns 385
DB 367 GCGGCGCCCTCGCTATTACTATTAACAAGCCGATTCGACAAAGC----- 413
QY 385 erSerMetTyrrLysTyrrProSerAspLeuProTyrrMetSerSerTyrrHisGlyLysArgT 405
DB 414 -----AAGGGAAACGGT 426
QY 405 yAlaTyrrLysPheAspPheHisGlylleAlaGlnAlaLeuGlnProHis----- 421
DB 427 TCACCTACAAAGTTCAATTTCAACAACAGCTGTGTCGCAATTACCCATTGATGATG 486
  
```









TELEFAX: (512) 474-7577  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2410 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 us-08-780-835B-1

## Alignment Scores:

Pred. No.: 2 42e-18 Length: 2410  
 Score: 279.50 Matches: 108  
 Percent Similarity: 35.22% Conservative: 41  
 Best Local Similarity: 25.53% Mismatches: 115  
 Query Match: 10.80% Indels: 160  
 Gaps: 15

US-09-902-772-4 (1-478) x US-08-780-835B-1 (1-2410)

```

QY 45 SerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnProProAlaArgVal 64
Db 634 TCAGAGAGGAAACCCAGAGTCCCC-----GCACAGACCCCGCCCTGTCCCT 678
QY 65 ThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAsp 83
Db 679 GCACAGAGAGGAAACCCAGAGTCCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
QY 84 -----AspCysSerValAlaLys 89
Db 739 TCGGCATCAAGTCCCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 798
QY 90 GlysLysMetValSerSerSerAspAsnValGlyMetAsnIleValProAla 109
Db 799 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
QY 110 GlnLysHisIleProProProAsnMetThrThrAsnGlnValArgValIleValProAla 129
Db 829 CCTCAGAGTCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 888
QY 130 AspProThrLeuThrSerThrAspHisValArgIleThrLeuGlnIleValValLysGlu 149
Db 889 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906
QY 150 TyrosylLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLysGluLeuCys 169
Db 907 -----TGT 909
QY 170 LysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAspIleLeuLeu 189
Db 910 -----GCCACTCCTCA 921
QY 190 SerHisLeuHisTyrLeuArgGluThrProLeuProHisIleThrSerAspAspValAsp 209
Db 922 -----CATCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 969
QY 210 LysAlaLeuGlnAsnSerProArgLeuMetHisAlaArgAsnThrGlyAlaThrPhe 229
Db 970 ACCAAGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1029
QY 230 IlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThrArgProAsp 249
Db 1030 -----ACCCCTGTAGCAAGAGCTGGCCAGC 1056
QY 250 LeuProTyrGlnGlnAlaArgArgSerAla-----Tyr 260
Db 1057 -----CCGCTT---CAAGCCAGAGGAGGAGTCAAGTGGGACACAGTACCCAGGGGAGGAGTGG 1110
QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 1111 TGATCAAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170
QY 281 LysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIle----- 294

```

```

Db 1171 TGTACTCCACCCAGAGGAGCTTCTGTGACCCCTCTCCAGGTGATGAGATGAGGTTATG 1230
QY 295 -----Leu 295
Db 1231 GCTATGAAAAATCCCTTCGACCATTCACAGATGATGCTGTGATGTCTCCCTAAAAAATTG 1290
QY 296 GlyProThrSerSerArg-----LeuAlaAsnProGly----- 306
Db 1291 AAGGAGACATCAACAGAGAGAGAGAGATGAGAGCTTCCGGAGAG-GGGCCACCTTACCAGCCG 1349
QY 307 SerLysGlnIleGlnIleThrPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 326
Db 1350 CGGGGTGCTTACACAGTGTGAGAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1409
QY 327 AsnCysIleThrTrpGluGlyThrAsnGlyLysIlePheLysMetThrAspProAspGluVal 346
Db 1410 CATTCATTCGTTGGACAGAGCGGGGAGATGAGATTAACTAATTTGAACCTGAAGAGGTT 1469
QY 347 AlaArgArgTrpGlyLysGlyLysSerLysProAsnMetAsnIleAspLysLeuSerArg 366
Db 1470 GCCAGGCTCTGGGTATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529
QY 367 AlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisProProGlnSerSer 386
Db 1530 TCGCTGGATACATATTTATGAGAAAGCATCATGCAAGAGTG----- 1571
QY 387 MetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisGlyLysArgTyrAla 406
Db 1572 -----GCTGGCAGAGCTTACGTCG 1589
QY 407 TyrLysPhe 409
Db 1590 TACAAATTT 1598

```

Search completed: July 28, 2003, 07:41:48  
 Job time : 79.3165 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 06:34:34 ; Search time 301.516 Seconds  
(Without alignments)  
3270.525 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588  
Sequence: 1 MASTIKALSVSDSLFE.....IYPTRLPAHMSHLGTY 478

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1439767 segs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-O=/cgn2\_1/USPTO.spool/US09902772/r/unat.23072003\_093658\_14962/app\_query.fasta.1.1294  
-DB=Published\_Applications\_NA -QFMT=fastap -SUFFIX=trnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09902772@cgn.1.1.82@unat.23072003\_093658\_14962  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published\_Applications\_NA:\*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
13:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	2436	94.1	1528	11 US-09-902-772-3

2	2262.5	87.4	1447	11	US-09-902-772-1	Sequence 1, Appl
3	2180.5	84.3	3166	15	US-10-205-823-98	Sequence 98, Appl
4	801	31.0	567	10	US-09-864-761-18410	Sequence 18410, A
5	716	27.7	473	10	US-09-864-761-20472	Sequence 20472, A
6	638	24.7	467	12	US-09-918-995-23356	Sequence 23356, A
7	565	21.8	549	10	US-09-923-779-54	Sequence 54, Appl
8	503	19.4	420	10	US-09-864-761-3705	Sequence 3705, Ap
9	503	19.4	454	10	US-09-864-761-1653	Sequence 1653, Ap
10	481	18.6	1752	10	US-09-850-799-1	Sequence 1, Appl
11	444	17.2	2188	11	US-09-920-3004-1716	Sequence 1716, Ap
12	444	17.2	2188	15	US-10-033-528-1716	Sequence 1716, Ap
13	444	17.2	2268	11	US-09-920-3004-1693	Sequence 1693, Ap
14	444	17.2	2268	15	US-10-033-528-1693	Sequence 1693, Ap
15	443.5	17.1	3692	15	US-10-106-698-824	Sequence 824, App
16	442.5	17.1	2269	11	US-09-954-531-955	Sequence 955, App
17	430.5	16.6	1884	11	US-09-925-300-420	Sequence 420, App
18	382	14.8	348	10	US-09-864-761-29364	Sequence 29364, A
19	382	14.8	477	10	US-09-864-761-5035	Sequence 5035, Ap
20	382	14.8	519	10	US-09-864-761-12798	Sequence 12798, A
21	377	14.6	2532	15	US-10-037-270-869	Sequence 869, App
22	374	14.5	553	10	US-09-864-761-7427	Sequence 7427, Ap
23	371	14.3	225	10	US-09-864-761-21761	Sequence 21761, A
24	311	12.0	473	12	US-09-918-995-27904	Sequence 27904, A
25	305	11.8	3178	15	US-10-108-605-124	Sequence 124, App
26	305	11.8	3178	15	US-10-108-605-128	Sequence 128, App
27	303	11.7	416	12	US-09-918-995-8472	Sequence 8472, App
28	299.5	11.6	2064	12	US-09-884-363-3	Sequence 128, App
29	299.5	11.6	2064	15	US-10-167-485-1	Sequence 3, Appl
30	299.5	11.6	2333	11	US-09-920-3004-1788	Sequence 1, Appl
31	299.5	11.6	2333	11	US-09-880-107-3316	Sequence 1788, Ap
32	299.5	11.6	2333	15	US-10-033-528-1788	Sequence 3316, Ap
33	297.5	11.5	1894	10	US-09-126-9458-1	Sequence 1788, Ap
34	297.5	11.5	1905	10	US-09-866-356-2	Sequence 2, Appl
35	297.5	11.5	1942	15	US-10-157-031-83	Sequence 83, Appl
36	297	11.5	4106	15	US-10-084-817-187	Sequence 187, App
37	280.5	10.8	1087	15	US-10-102-806-215	Sequence 215, App
38	280.5	10.8	2952	10	US-09-925-297-257	Sequence 257, App
39	279.5	10.8	2410	12	US-09-884-363-1	Sequence 1, Appl
40	279.5	10.8	165	10	US-09-864-761-23294	Sequence 23294, A
41	278.5	10.8	1894	10	US-09-841-9634-1	Sequence 1, Appl
42	276	10.7	300	15	US-10-177-063-15	Sequence 15, Appl
43	267	10.3	2256	15	US-10-098-841-317	Sequence 317, App
44	261	10.1	300	15	US-10-177-063-16	Sequence 16, Appl
45	244.5	9.4	1426	10	US-09-925-297-309	Sequence 309, App

#### ALIGNMENTS

RESULT 1  
US-09-902-772-3  
Sequence 3, Application US/09902772  
Patent No. US20020164739A1  
GENERAL INFORMATION:  
APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et  
TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
TITLE OF INVENTION: the Proteins  
FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001  
CURRENT APPLICATION NUMBER: US/09/902.772  
PRIOR FILING DATE: 1997-06-18  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 3  
LENGTH: 1528  
TYPE: DNA  
ORGANISM: c-ery gene, chicken DNA  
US-09-902-772-3  
Alignment Scores:  
Pred. No.: 4,246-254  
Score: 2436.00  
Percent Similarity: 91.97%  
Length: 1528  
Matches: 458  
Conservative: 0

Best Local Similarity: 91.97% Mismatches: 0  
 Query Match: 94.13% Indels: 40  
 DB: 11 Gaps: 2

US-09-902-772-4 (1-478) x US-09-902-772-3 (1-1528)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValIValSerGluAspGlnSerLeuPheGlu 20  
 DB 63 ATGCAAGCACTATTAAAGAGCATTTTCAGTGGTGAAGCAACAGCTCTGTTGAG 122  
 QY 21 CysAlaIleTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 DB 123 TGTGCTACGAGATGCCCACTTGCAAGACAGAAATGACAGGCTCTCTCCAGTGA 182  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60  
 DB 183 TATGGCAACATCAAAAGATGAGCCCGGGTCCAGAGAGCACTGGTATACACAGCC 242  
 QY 61 ProAlaArgValIleThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 DB 243 CCGGCCAGAGTTACCATTAAGATGAGTGAACCCAAACAGGTTAATGGGTCAAGGAA 302  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnVal 100  
 DB 303 TCACCTATGACTGACCGTGGCAAGAGAGGAAATGGTTACAGCTCAGACAAATGTT 362  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThr 120  
 DB 363 GGGATGAACATATGGAACCTACATGAGAGAGACATATTCGGCTCAAAATATGACAAAC 422  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuThrIleuThrAspHisValArg 140  
 DB 423 AATGAACGAGAGATTATGTCAGCAGATCTACGTTATGAGACACAGACCATGTACG 482  
 QY 141 GlnTrpLeuGluTrpAlaValLysGlyLysLeuProAspValAspIleLeuPhe 160  
 DB 483 CAGTGGCTGAGTGGAGAGTGAAGATATGTTCTTCACAGCGTGGACATCTTGTTC 542  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 DB 543 CAGAACATGTATGGGAAGAGTGTGTAATAATGACCAAGATGACTTCCAGAGACTCAG 602  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200  
 DB 603 CCGAGCTATACGAGATATCTCTCTGTCACACTACACCTCAACCTCAAGAGACTCTCT 662  
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220  
 DB 663 CCACATTTGACTCAGATGATGTTGATAGAGCCCTTACAAACTCTCCACGGTTATGAT 722  
 QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240  
 DB 723 GCTAGAACAACAGAGAGGACCTTTATTTTCCAAATACATCAGATTACCCGAAAGCA 782  
 QY 241 ThrGlnArgGlnLeuThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 260  
 DB 783 ACCCAAGATTAACAACAGGCAAGGATTTACTTATGAGCAAGAGAGATCAGCGTGG 842  
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280  
 DB 843 ACGAGTCACAGCATCCCACTCAGTCAAAAGCTAACCCACATCATCTTCAACAGTGGCC 902  
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300  
 DB 903 AAAACAGAAAGACAGCGTCTCTCAGTTAGATCTTATCAGATTCCTGGACCGACAGAGC 962  
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 320  
 DB 963 CGCTTGTCAATCCAGGAGTGGGAGATACAGTATGAGCAAGTCCACTACGAGCTTCTG 1022  
 QY 321 SerAspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyLysIleLysMet 340  
 DB 1023 TCGGACAGCTCCAACTCCAACTGATCCTGGGAGGGACAAATGGGGAGTTCAGATG 1082

QY 341 ThrAspProAspGluValAlaArgArgTrpGlyLysSerLysSerLysProAsnMetAsn 360  
 DB 1083 ACAGACCCCTGATGAAGTGGCTCGGCGTTGGGGAGAGAGAAAGCAAACTAATCATGAA 1142  
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380  
 DB 1143 TATGACAACTCAGCCCTGACCTTCCTACTACTATGACAAATAATATATGACTAAAGTT 1202  
 QY 380 ----- 380  
 DB 1203 CATGCTAAACGCTATGCTTCAAAATTTGATTTCCACGGAATCGCTACGGCCCTCCAGCT 1262  
 QY 381 HisProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400  
 DB 1263 CACCCCTCAGATCATCATGATACAAATATACCATCAGACCTCCCTCATGATGAGTCTTAC 1322  
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisIleAlaGlnAlaLeuGlnPro 420  
 DB 1322 ----- 1322  
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440  
 DB 1323 CATGCAACCCCGAAGAGATGAACTTTGTAGCTCCCATGCCCTGCTTGGCCGTAAAC 1382  
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 460  
 DB 1383 TCATCCAGCTTTTGTGCTGCCCTTAATCATCTAGTAATGACCAACTGAGAGCATCTAC 1442  
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478  
 DB 1443 CCCAATACAGGCTGCCAGCTGCTCATATGCTTCCATCTTGGCACCTACTAC 1496

# RESULT 2

US-09-902-772-1  
 : Sequence 1, Application US/09902772  
 : Patent No. US20020164739A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et  
 : TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
 : TITLE OF INVENTION: the Proteins  
 : FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001  
 : CURRENT APPLICATION NUMBER: US/09/902,772  
 : PRIOR FILING DATE: 2001-07-12  
 : PRIOR APPLICATION NUMBER: US/08/878,177  
 : NUMBER OF SEQ ID NOS: 7  
 : SOFTWARE: Patentin Ver. 2.0  
 : SEQ ID NO 1  
 : LENGTH: 1447  
 : TYPE: DNA  
 : ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA  
 : US-09-902-772-1

## Alignment Scores:

Pred. No.: 2,53e-235 Length: 1447  
 Score: 2262.50 Matches: 430  
 Percent Similarity: 86.35% Conservative: 1  
 Best Local Similarity: 86.35% Mismatches: 1  
 Query Match: 87.42% Indels: 67  
 DB: 11 Gaps: 3

US-09-902-772-4 (1-478) x US-09-902-772-1 (1-1447)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValIValSerGluAspGlnSerLeuPheGlu 20  
 DB 63 ATGCAAGCACTATTAAAGAGCATTTATCAGTGTAGTGAACACAGTCTTGTGAG 122  
 QY 21 CysAlaIleTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 DB 123 TGTGCTACGAGATGCCCACTTGCAAGAGAGAAATGACACACCTCTTCCAGTGA 182  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60

```

183 TATGGCAACATCAAAATAGAGCCGGCGCTTCCCGACGAGCTGTTATCAGAGCC 242
Db
61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db
243 CCGGCGAGAGTACCATTAAAGATGAGTGAACCAACCAAGTAAAGGGTCAAGCAAT 302
Db
81 SerProAspPaspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db
303 TCACCTGATGACTCAGCGTGGCAAAAGAGGAAAGTGTACAGTTCAGACAAATGTT 362
Db
101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120-
Db
363 GGGATGAACCTATGAGACGTACATGAGAGAGAGCATATTCGGCTCCAAATATGACAACC 422
Db
121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrPserThrAspHisValArg 140
Db
423 AATGACGAGAGAGTATGTTGTCAGAGATCTCAGTATGAGACACAGACCATGTACGG 482
Db
141 GlnThrLeuGluTyrPAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
Db
483 CAGTGGCTGAGTGGGAGTGAAGAGATGATGCTTCCAGACGTGAGACATCTTGTGTTC 542
Db
161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db
543 CAGAACATTTGATGGGAAAGAGTGTGTAATATGACCAAGATGACTTCCAGAGACTCAGC 602
Db
181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLysHisTyrLeuArgGluThrProLeu 200
Db
603 CCGAGCTATACGAGATATCTCTGTCACACCTACCTACCTACCTACCTACCTACCTACCT 656
Db
201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
Db
656 ----- 656
Db
221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
Db
657 -----GGAGCCACTTTATTTTCCAAATACATCAGTTTACCGAAGCA 701
Db
241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrp 260
Db
702 ACGGAAGAAATAAACAAAGCCAGATTTACCTTATGAGCAAGGAGAGATCAGCGTGG 761
Db
261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db
762 ACGAGTACAGCCATCCCACTCAGTCAAAAGCTACCAACCATATCTTCAACAGTGGCC 821
Db
281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db
822 AAAACAGAAAGCAGCGCTCCATGATGATCTTATCAGATTTCTTGACCGACGACGACG 881
Db
301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuLeuLeu 320
Db
882 CGTCTTGAAATCCAGGAGGTGGGAGATACAGCTATGCGAGTTCCTACTGAGCTTCTG 941
Db
321 SerAspSerSerAsnSerAsnCysIleThrTyrGluGlnThrAsnGlyGluPheLysMet 340
Db
942 TCGGACAGCTCAACTCAACTGATCACCCTGGGAGGGGCAAAATGGGAGTTCAAGATG 1001
Db
341 ThrAspProAspGluValAlaLysArgTyrPAlaValLysGlyLysSerLysProAsnMetAsn 360
Db
1002 ACAACACCTGATGAAGTGGCTCGGCGTTGGGAGAGAGAGAAACCAACCTTAACATGAC 1061
Db
361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db
1062 TATGACAAACACACCGCTGCTACTTCTACTACTATGACAAATAATATATGACTAAAGTT 1121
Db
380 ----- 380
Db
1122 CATGTAACCGCTATGCTACAAATTTGATTTCACGGAATCGCTCAGGCCCTCAGCGCT 1181
Db
381 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400

```

```

Db
1182 CACCTCCAGATATCATCATGTAACAATACCACATCAGACCTCCCTACATGATGTTCTAC 1241
Db
401 HisGlyLysArgTyrAlaTyrLysPheAspHisGlyIleAlaGlnAlaLeuGlnPro 420
Db
1241 ----- 1241
Db
421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db
1242 CATGACACACCCCGAAGAGATGACTTGTAGCTCCCATTCCTCCGCTTGGCCGTAAAC 1301
Db
441 SerSerSerPhePheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyLysTyr 460
Db
1302 TCATCCACACTTTTTCGTCGCCCTTATATCATAGTGAATTCACCAACGAGGAGCATTCAC 1361
Db
461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db
1362 CCATATACAGAGCTGCCAGCTGCATATGCTTCCATCTTGGCAGCCTACTAC 1415
Db
1362 ----- 1415

RESULT 3
US-10-205-823-98
; Sequence 98, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatz, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-98

Alignment Scores:
Pred. No.: 6,13e-226 Length: 3166
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 15 Gaps: 4

US-09-902-772-4 (1-478) x US-10-205-823-98 (1-3166)
Db
2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
Db
281 GCAGCTCATATCAAGAGAGCCCTTATGATGAGTGAAGAGACAGTGTGTTGAGAGTGT 340
Db
22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41

```

```

Db 341 GCCACGGAACGACACCTGGCTAAGACAGAGATGACGCGCTCCTCCAGGACAT 400
Qy 42 GYGLNINrSerLysMetSerProArValProGlnGlnAspTrpLeuSerG1ProPro 61
Db 401 GGACAGACTTCCAAAGATGAGCCCGCTCCCTCAGCAGGATGGCTTCTCAACCCCA 460
Qy 62 AlaArValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 461 GCCAGGGTCACCATCAAAATGGAATGTAACTCCAGCCAGTGAATGGCTCAAGAACTCT 520
Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspValGly 101
Db 521 CCTATGAAATCACTGTGGCCAAAGCGGGAAGATGGGGCAGCCCAACCCGCTTGGG 580
Qy 102 MetAsnTrpGlySerTrpMetGluGluLysHisIleProProAsnMetThrAsn 121
Db 581 ATGAACrTACGCGACCTACATGAGAGAGACATGCCACCCCAACATACCAACGAGAC 640
Qy 122 GlnArGArValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db 641 GAGCGCAGAGTATTCGTGCCAGAGATCCTACGCTATGAGTACAGACCATGTGGCGAG 700
Qy 142 TrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPheGln 161
Db 701 TGGCTGAGAGTGGGGGGAAGAAATATGCGCTTCAGACGTCACATCTTGTATTCAG 760
Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 761 AACATCGATGGAGAGAACTGTGCAAGATGACCAAGACGACCTTCAGAGGCTCAACCCC 820
Qy 182 SerTrpAsnAlaAspIleLeuLeuSerHisLeuHisTrpLeuArgGluTrpProLeuPro 201
Db 821 AGCTACACAGCGCGACATCTCTCTCATCTCCACTACCTCAAGAGACCTCTCTCA 880
Qy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProAlaGlyLeuMetHisAla 221
Db 881 CATTTGACTTCAGATGATGTGATTAAGCTTACAAAACCTCCACGGTATATCAGTCT 940
Qy 222 ArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyProGluAlaThr 241
Db 941 AGAAGACAC----- 949
Qy 242 GlnArGleThrThrArgProAspLeuProTrpGluGlnAlaArgArgSerAlaTrpThr 261
Db 950 -----GATTACCATATGAGCCCGCCAGAGATGAGCTCGAGCC 988
Qy 262 SerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 989 GGTACGCGCCACCCCGCCCGCCAGTCAAAAGCTCTCAACCATCTCTCCACAGTGGCC 1048
Qy 281 LysThrGlnAspGlnArgProGlnLeuAspProTrpGlnIleLeuGlyProThrSerSer 300
Db 1049 AAAACTGAAGACCGCGCTCCCACTAGATTCCTTATCAGATCTTGGACCAACAAGTAGC 1108
Qy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeuLeu 320
Db 1109 CGCCTTCCAATCCAGCAGTGGCCAGATCCAGCTTGGCGAGTTCTCTCTGAGACTCTCG 1168
Qy 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlnGlyThrAsnGlyGlnPheLysMet 340
Db 1169 TCGGACAGCTTCCAACTCAGCTGATCACTCGGGAAGGACCAACGAGGAGTTAAAGATG 1228
Qy 341 ThrAspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsn 360
Db 1229 ACGGATCCCGACGAGGAGTGGCCGCTGGGGAGAGGGAAGACCAACCAACATATGAGC 1288
Qy 361 TyrAspLysLeuSerArgAlaLeuArgTrpTyrTrpTyrAspLysAsnIleMetThrLysVal 380
Db 1289 TACGATTAAGCTACAGCCCGCTCTCTTACTATATACAAAGACATCATATGACCAAGGTC 1348
Qy 381 HisProProGluSerSerMetTrpLysTyProSerAspLeuProTyTrpMetSerSerTy 400

```

```

Db 1348 ----- 1348
Qy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1349 CATGGGAAGCCCTACGCTTCAAGATTCGACTTCCACGGAGATGCGCCAGCCCTCAGGCC 1408
Qy 420 ----- 420
Db 1409 CACCCCGGAGTATCTCTGTACAGTACCCCTCAGACCTCCGTCATATGGCTCTAT 1468
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1469 CACGCCACCCACGAGAAGATGAATCTTGTGGCGCCACCTCCAGCCCTCCCGTACCA 1528
Qy 441 SerSerSerPhePheAlaAlaProAsnProTyTrpAsnSerProThrGlyGlyIleTyr 460
Db 1529 TCTTCCAGTTTTTTTGTGTCGCCCAACCCATACGTAATTCACCACTGGGGGTATATAC 1588
Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyTrp 478
Db 1589 CCCAACACTAGCTCCCGACAGCCATATGCTTCTCATCTGCGGCACTTACTAC 1642

RESULT 4
US-09-864-761-18410/C
; Sequence 18410, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117

```

```

? SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 18410
? LENGTH: 567
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AP000163.1
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
? OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
? OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.7
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
? OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
? OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 9.00e-99
? OTHER INFORMATION: EST_HUMAN HIT: R87572.1, EVALUE 0.00e+00
? OTHER INFORMATION: NT HIT: M17254.1, EVALUE 0.00e+00
? OS-09-864 -761-18410

```

Alignment Scores:		
Pred. No.:	1,8e-77	Length: 567
Score:	801.00	Matches: 151
Percent Similarity:	78.17%	Conservative: 3
Best local Similarity:	76.65%	Mismatches: 40
Query Match:	30.95%	Indels: 2
DB:	10	Gaps: 2

US-09-902-772-4 (1-478) x US-09-864-761-18410 (1-567)

OY		302	IeAlaHisProGlnSerIylInIleGlnLeuTrpGlnPheLeuGlnIleLeuSer      :::	332
Db		534	CfTAcCTCATTAGACAGTGGCCAGATTCACGCTTTGGACTCTCTCGAGACTCTCTGC 	475
OY		322	AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyIuPheLysMetThr      :::	344
Db		474	GAGAGCTCAACTCCACGTGCACTCACCTGGGAAGCACCACCGGAGGATTCAGATGACG 	415
OY		342	AspProAspGluValAlaArgArgTrpGlyAlaGlySerLysProAsnMetAsnTrp 	361
Db		414	GATCCCGAGAGGTGGCCGGCGGGGAGAAGCGAAGAACCAACCATCACTAACTAC 	355
OY		362	AspLysIleuSerArgAlaLeuArgTrpGlyTrpTyrAspLysAsnIleMetThrLysValHis 	381
Db		354	GATAAGCTAGGCCGGGCCCTCCGTTACTACTATGACAGAACATCATAGCCAAGATC--- 	298
OY		382	ProProGluSerSerMetLysTrpProSerAspLeuProTyrlMetSerSerTrpHis 	401
Db		297	-----CAT 295	
OY		402	GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 	420
Db		294	GGGAAGCGCTACGCCCTACAAAGTTCACACTTCACACGGAGATGGCCAGGCCCTCACCCCCAC, 	235
OY		421	-----His 421	
Db		234	CCCCGGAGTCACTCTGTACAAAGTACCCCCTCAGACCTCCCGTACATGGGCTCTTAATAC 	175
OY		422	AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 	441
Db		174	GCCACCCACAGAGAATGAATTTGTGGCGGCCACCCCTCCAGGCCCTCCCGTGACATCT 	115
OY		442	SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 	461
Db		114	TCCAGTTTTTTTGGCTGCCCAAACCCAACTACGGAATTCACCAACTGGGGGATATATACCCC 	55
OY		462	AsnThrArgLeuProAlaAlaHisMetProSerHisIleuGlyLysTrpTyrTrpTyr 478      :::	
Db		54	AACACTAGGCTCCCCACACGACCATATGCTTCTCATCTGGGCACTTACTAC 4 	

US-09-864-761-20472/c  
; Sequence 20472, Application US/09864761  
; Patent No. US20020048763A1  
; General Information

: APPLICANT: Penn, Sharon G.  
 :  
 : APPLICANT: Rank, David R.  
 :  
 : APPLICANT: Hanzel, David K.  
 :  
 : APPLICANT: Chen, Wensheng  
 :  
 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR  
 :  
 : TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

;  
PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR APPLICATION NUMBER: IIS 60/236.359

PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: BCT/US01/006666

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30  
; PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: 2001/000654

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

; PRIOR FILING DATE: 2001-01-30

; PRIOR FILING DATE: 2001-01-30

;  
PRIOR FILING DATE: 2001-01-30

;; PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2000-09-21

PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/114,203  
 ; PRIOR FILING DATE: 2001-01-29

```

; NUMBER OF SEQ ID NOS: 49117
SOFTWARE: ARCGEN 1.4.1

```

; SEQ ID NO 20472  
TTCGCTT 473

TYPE: DNA

```

;
FEATURE:

```

OTHER INFORMATION: EXPRESSED IN BT474, S

OTHER INFORMATION: EXPRESSED IN FETAL LI

OTHER INFORMATION: EXPRESSED IN BONE MARROW

OTHER INFORMATION: EXPRESSED IN HBL100,

OTHER INFORMATION: EST\_HUMAN HIT: R87572

OTHER INFORMATION: NT HIT: M17254.1, EVAR

7 / 2 2 7  
T C /  
C C C  
C C C  
C C C

Alignment scores:	
Pred No.:	2 27e-68
Length:	

score:	16.00	match
--------	-------	-------

Percent Similarity: 77.40%  
 Best Local Similarity: 76.84%  
 Query Match: 27.67%  
 DB: 10 Gaps: 2

US-09-902-772-4 (1-478) x US-09-864-761-20472 (1-473)

Qy 306 GlyserglycInleuInleuTrpGlnPheLeuLeuGlnLeuLeuSerSerSerSer 325  
 Db 473 GGAGTGGCCAGATCCAGCTTGGCAGTCTCTGGAGCTCCGTGGAGAGAGCTCCAC 414  
 Qy 326 SerancysIleHrTrpGlnGlyThrAsnGlyLupheLysMetThrAspProAspGlu 345  
 Db 413 TCCAGCTGCATCACCTGGGAGGACCAACCGGGAGTTCAAGATGACGATCCGACAG 354  
 Qy 346 ValAlaArGArGTpGlyGluArgLysSerLysProAsnMetAsnTyraAspLysLeuSer 365  
 Db 353 GTGGCCCGCCCTGGGAGAGCGAGCAAGCAACCAACATGAACTGATTAACCTCAGC 294  
 Qy 366 ArgAlaLeuArGTyTyrTyraAspLysAsnIleMetThrLysValHisProProGluSer 385  
 Db 293 CGGCCCTCGTACTACTATGACAGACATCATGACCAAGTGC----- 249  
 Qy 386 SerMetTyLysTyTyrProSerAspLeuProTyMetSerSerTyTHisGlyLysArgTyr 405  
 Db 248 -----CATGGGAGACGCTAC 234  
 Qy 406 AlaTyrlLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro----- 420  
 Db 233 GCCTACAGTTCCGACTTCACGCGGATCGCCAGGCCCTCCAGCCCGGAGTCA 174  
 Qy 421 -----HisAlaHisProGln 425  
 Db 173 TCTCTGACAGTACCCCTCAGACCTCCGTCATGATGAGGCTCTTACAGGCCACACAG 114  
 Qy 426 LysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhe 445  
 Db 113 AAGATGAACTTGTGGGCGCCACCCCTCCAGCCCTCCGAGCATCTCCAGATTCTTTT 54  
 Qy 446 AlaAlaProAsnProTyTyrPAsnSerProThrGlyLysIleTyProAsn 462  
 Db 53 GCTGCCCAACCCATCTGAAATTCACAACTGGGGGTATATATACCCCAAC 3

# RESULT 6

US-09-918-995-23356  
 : Sequence 23356, Application US/0918995  
 : Publication No. US20030073623A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Hyseq, Inc.  
 : TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
 : FILE REFERENCE: 20411-756  
 : CURRENT APPLICATION NUMBER: US/09/918,995  
 : PRIOR FILING DATE: 2001-07-30  
 : PRIOR APPLICATION NUMBER: US/09/235,076  
 : NUMBER OF SEQ ID NOS: 38054  
 : SOFTWARE: FastSeq for Windows Version 3.0  
 : SEQ ID NO 23356  
 : LENGTH: 467  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : NAME/KEY: misc\_feature  
 : LOCATION: (1)..(467)  
 : OTHER INFORMATION: n = A,T,C or G  
 : US-09-918-995-23356

## Alignment Scores:

Pred. No.: 6,36e-60 Length: 467  
 Score: 638.00 Matches: 120  
 Percent Similarity: 97.66% Conservative: 5  
 Best Local Similarity: 93.75% Mismatches: 3

Query Match: 24.65%  
 DB: 12 Gaps: 0

US-09-902-772-4 (1-478) x US-09-918-995-23356 (1-467)

Qy 1 MetaLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 84 ATGGCCAGCACTATTAAAGAAAGCCTTATCACTTGTGATGAGCAGCAGTGGTTGAG 143  
 Qy 21 CysAlaTyrgLysSerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGlu 40  
 Db 144 TGTGCTTACGGAGCAACCACTGTGGCTTAAGACAGATGACCGCTCTCTCCAGGAC 203  
 Qy 41 TyrgLysInThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 Db 204 TATGACAGACTTCCAAAGATGAGCCACGCGCTCCCTCAGCAGATTTGGTGTCAACCC 263  
 Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 264 CCACCCAGGGTCAACCATCAAAATGAAATGTAACCTGACCGAGTGAATGGCTCAAGGAC 323  
 Qy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 Db 324 TCTCTGATGATGACAGTGTGGCCAAAGCGGGAAGTGGTGGCAGCCACAGACAGCTT 383  
 Qy 101 GlyMetAsnTyrgLysSerTyMetGluGluLysHisIleProProProAsnMetThr 120  
 Db 384 GGGATGAACTAGGCGCAGCTACATGAGAGAGACATGACCCCAACATATGACACAG 443  
 Qy 121 AsnGluArgArgValIleValPro 128  
 Db 444 AACGAGCGCAGAGTTATCTGCCA 467

# RESULT 7

US-09-923-779-54  
 : Sequence 54, Application US/09923779  
 : Patent No. US20020076721A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Pyle, Ruth A.  
 : APPLICANT: Kalos, Michael D.  
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 : FILE REFERENCE: 210121.553  
 : CURRENT APPLICATION NUMBER: US/09/923,779  
 : PRIOR FILING DATE: 2001-08-06  
 : NUMBER OF SEQ ID NOS: 155  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 54  
 : LENGTH: 549  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : NAME/KEY: misc\_feature  
 : LOCATION: 356, 363, 417, 433, 442, 461, 463, 464, 469, 479, 485, 489,  
 : LOCATION: 537, 545  
 : OTHER INFORMATION: n = A,T,C or G  
 : US-09-923-779-54

## Alignment Scores:

Pred. No.: 6.61e-52 Length: 549  
 Score: 565.00 Matches: 110  
 Percent Similarity: 71.15% Conservative: 1  
 Best Local Similarity: 70.51% Mismatches: 4  
 Query Match: 21.83% Indels: 41  
 DB: 10 Gaps: 2

US-09-902-772-4 (1-478) x US-09-923-779-54 (1-549)

Qy 344 AspGluValAlaArgArGTpGlyGluArgLysSerLysProAsnMetAsnTyraAspLys 363  
 Db 2 GACGAGTGGCCCGCGCTGGGAGAGCGGAGAGCAACCAACATGAACTGACGATAG 61

QY 364 LeuSerArgAlaLeuAlaArgTyrTyrTyrAspLysAsnIleMetThrLysValHisProPro 363  
 |||||||  
 Db CTCACCCGGCGCCCTCGTTACTACTATGACAGAAATCATGATGACCAAGGTC----- 112  
 QY 384 GluSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisGlyLys 403  
 |||||||  
 Db 113 -----CATGGGAG 121  
 QY 404 ArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro----- 420  
 |||||||  
 Db 122 CGCTACGGCTACAAAGTTCGACTTCCACGGGAGTCCGCCACGGCCCTCCAGCCCAACCCCCG 181  
 QY 421 -----HisAlaHis 423  
 Db 182 GAGTCATCTCTGTACAGATACCCCTCAGACCTCCCTACATGGGCTCTATGACGGCCAC 241  
 QY 424 ProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSer 443  
 |||||||  
 Db 242 CCACGAAAGATAACTTTGTGGCGCCGCCACCTCCAGCCCTCCCGGATCTTCAGT 301  
 QY 444 PhePheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyLysTyrProAsnThr 463  
 |||||||  
 Db 302 TTTTGTGGTGGCCCAACCATATGGAATTACCAACATGGGGGTATATACCCCNCACT 361  
 QY 464 ArgLeuProAlaAlaHisMetProSerHisLeu-GlyThrTyrTyr 478  
 |||||||  
 Db 362 ANGCTCCCAACGACCATATGCTTTTCATCTGGGGGCACTTACTAC 407  
 RESULT 8  
 US-09-864-761-3705/C  
 : Sequence 3705, Application US/09864761  
 : Patent No. US2002048765A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Penn, Sharon G.  
 : APPLICANT: Rank, David R.  
 : APPLICANT: Hanzel, David K.  
 : APPLICANT: Chen, Wensheng  
 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 : FILE REFERENCE: Aecmics-X-1  
 : CURRENT APPLICATION NUMBER: US/09/864,761  
 : CURRENT FILING DATE: 2001-05-23  
 : PRIOR APPLICATION NUMBER: US 60/180,312  
 : PRIOR FILING DATE: 2000-02-04  
 : PRIOR APPLICATION NUMBER: US 60/207,456  
 : PRIOR FILING DATE: 2000-05-26  
 : PRIOR APPLICATION NUMBER: US 09/632,366  
 : PRIOR FILING DATE: 2000-08-03  
 : PRIOR APPLICATION NUMBER: GB 24263.6  
 : PRIOR FILING DATE: 2000-10-04  
 : PRIOR APPLICATION NUMBER: US 60/236,359  
 : PRIOR FILING DATE: 2000-09-27  
 : PRIOR APPLICATION NUMBER: PCT/US01/00666  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00667  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00664  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00669  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00665  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00668  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00663  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00662  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00661  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00670  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: US 60/234,687  
 : PRIOR FILING DATE: 2001-01-30

```

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3705
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP00021.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
US-09-864-761-3705

Alignment Scores:
Pred. No.: 2,36e-45 Length: 420
Score: 503.00 Matches: 96
Percent Similarity: 70.29% Conservative: 1
Best Local Similarity: 69.57% Mismatches: 1
Query Match: 19.44% Indels: 40
DB: 10 Gaps: 2

US-09-902-772-4 (1-478) x US-09-864-761-3705 (1-420)
QY 361 TYRASPLYLSLEUSERATGALALEUARGTYRTYTYRASPLYLSASNIEMETHRLYSVAL 380
|||||
DB 418 TACGATAGCTCAGCGCGGCCCTCCGTTACTACTATGACAAAGAACATCATGACCAAGTC 359
|||||
QY 381 HISFPROGLJUSERSETMETTYRLYSTYTPROSERASPLEUPROTYRMEUSERSETYR 400
|||||
DB 359 ----- 359
QY 401 HISGLYLSATGTYRATYRLATYRLSPHEASPHENISGLYILEAGINALEUINPRO 420
|||||
DB 358 CATGGGAGCGCTACGCGCTACAAAGTTGGACTTCACGGGATCGCCAGGCCCTCCAGCCC 299
|||||
QY 420 ----- 420
DB 298 CACCCCGCGGAGTATCTCTTACAGTACCCCTCAACACCTCCCGTACATGGGCTCTAT 239
|||||
QY 421 HISALAHISPROGLINLYSMETASNPHENIALEAPROHISPROPROALALEUPROVALTHR 440
|||||
DB 238 CAGCGCCGACCCAGAGAAGATGAACTTGTGGCGGCCACCCCTCCAGGCCCTCCCGTGACA 179
|||||
QY 441 SERSESRPHEPHEALALAPROASNPROTYRTPASNSERPROTHRGILGYLIETYR 460
|||||
DB 178 TCTTCAGATTGTTTGTGTGGCCCCCAACCCATACGGAATTCCACCACTGGGGGTATATAC 119
|||||
QY 461 PROASNTPHARGLEUPROALAHISMEPROSERHISLEUGLYNTHRYTYR 478
|||||
DB 118 CCCAACACTAGGCTCCCAACAGCCATATGCTTCTATCTGAGGCACTTACTAC 65
|||||

RESULT 9
US-09-864-761-1653/C
; Sequence 1653, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1

```

CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
SEQ ID NO 1653  
LENGTH: 454  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AP000163.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95  
US-09-864-761-1653  
Alignment Scores:  
Pred. No.: 2,64e-45 Length: 454  
Score: 503.00 Matches: 96  
Percent Similarity: 70.29% Conservative: 1  
Best Local Similarity: 69.57% Mismatches: 1  
Query Match: 19.44% Indels: 40  
DB: 10 Gaps: 2  
US-09-902-772-4 (1-478) x US-09-864-761-1653 (1-454)  
QY 361 TYNASPLYSLEUSERATGALALEUARGTYRTYTYRSPLYSASNIEMETHLYSVAL 380  
|||||  
DB 452 TAGCATAGCTACACCGCGCCCTCCGTACTACTACATGACAAACATCATGACCAAGTC 393  
QY 381 HISPROGLUSERSERMETTYRLYSYRPROSERASPLEUPROTYRMETSERSEXTYR 400

DB 393 ----- 393  
QY 401 HISGLYLSARGTYRALATYRGLYSRHEASPHENISGLYILEAGINALALEUGLAPRO 420  
|||||  
DB 392 CATGGAGAGCGCTACGCGCTACAGTTCGACTTCCACGGAGTCCAGGCCCTCCAGCCC 333  
QY 420 ----- 420  
DB 332 CACCCCGGAGTATCTCTACAGTACCCCTCAGACCTCCCGTACATGGCTCTAT 273  
QY 421 HISALAHISPROGLINLYSMETASNPHEVALAIAHROHISPROFOLALEUPROVALTHR 440  
|||||  
DB 272 CACGCCACCCACAGAGATGAACTTGTGGCGCCACCTCCAGCCCTCCCGTGACA 213  
QY 441 SERISEREPHEALIAIAHROHISPROFOLALEUPROVALTHR 460  
|||||  
DB 212 TCTTCAGTTTTTTTGTGCTGCCCCAACCCTACTGGAATTCACCAACTGGGGGTATATAC 153  
QY 461 PROASNTHARGLEUPROALIAIAHISMETPROSERHISLEUGLYTHRTYR 478  
|||||  
DB 152 CCCAAGCTAGCTGCCGCCACGACGATATGCTTCTCATCTGGGCACTTACTAC 99  
RESULT 10  
US-09-850-799-1  
Sequence 1, Application US/09850799  
Patent No. US20020090647A1  
GENERAL INFORMATION:  
APPLICANT: Denieris, Eyan S.  
APPLICANT: Eyodoto, Dmitry V.  
APPLICANT: Hendricks, Timothy J.  
TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds  
TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases  
FILE REFERENCE: CASE-03898  
CURRENT APPLICATION NUMBER: US/09/850,799  
CURRENT FILING DATE: 2001-05-08  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/360,779  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-26  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 1752  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (112)..(1131)  
US-09-850-799-1  
Alignment Scores:  
Pred. No.: 4.57e-42 Length: 1752  
Score: 481.00 Matches: 122  
Percent Similarity: 44.73% Conservative: 18  
Best Local Similarity: 38.98% Mismatches: 68  
Query Match: 18.59% Indels: 105  
DB: 10 Gaps: 10  
US-09-902-772-4 (1-478) x US-09-850-799-1 (1-1752)  
QY 224 THRGLYGLYALATRPHEIIEPHEPROASNTHSERYALTYRPROGLIALATHRGINARG 243  
|||||  
DB 333 ACGGCGGGATGCGCACCCCAAGCTGACCTGCCCCCTCGGCTCCGCC 392  
QY 244 ILEHTRHTRPRPOASPLEUPROTYRGLUGINALIARGSERALATRPHTSERHIS 263  
|||||  
DB 393 ATCTCCACCGCCCA ----- GTCGCCGCC 416  
QY 264 SERHISPROTHRGINSERLYSALATHRGINPROSESERSETHRVALPROLYSTHGLU 283  
|||||  
DB 417 AGCGATGACAGACAGCGGACACCTCCAGCCCTCTGATACATGACTACAGATTC 476  
QY 284 ASPLINARG ----- PROGLINLEUASPROTYRGLINLEU 295  
|||||



```

Db      477 CGTCGAGATGCTCTTTTAAAGAGAGAGAGCCGAGCTGGGGCCG-----CT 527
QY      295 uGlyProThrSerSerArgLeuAlaAnProGlySerGlyIleIleIleuThrProIle 315
Db      528 GAGCCCTGGCGTACAGAAA-----GGCAGCGGGCAGATTCAGTTGTGGCGTT 575
QY      315 eLeuLeuGluLeuLeuSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 335
Db      576 TCtACTGAGACTGCTGGCAGACCGCGAGACCGCGGCTGCATCGCTGGAGAGGGCGCCA 635
QY      335 nGlyGluPheLysMetThrAspProAspGluValAlaArgArgTyrPglYgluArgLysSe 355
Db      636 CGCGAGTTCAGACTCAGCAGCCCGCAGAGGTGGCGCGAGCGCTGGGGGAGGCAAGAG 695
QY      355 rLysProAsnMetAsnTrpAspLysLeuSerArgAlaLeuArgTyrTyrTrpAspLysAs 375
Db      696 CAAGCCCAATATGAACTACAGCAAGCTAAGTCGAGCTCGCTACTACTACGCAAAA 755
QY      375 nIleMetThrLysValHisProProGluSerSerMetTyrLysTyrProSerAspLeuPr 395
Db      756 CATCATGACGAGGTG----- 771
QY      395 oTyrMetSerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAl 415
Db      772 -----CACGCAAGCGCTACGCTACCGCTTGACTTGACTTCAGAGGCGCTGGC 815
QY      415 aGlnAlaLeuGlnPro-----HisAlaHisPro----- 424
Db      816 ACAGGCTTGGCCAGCCACCCCGCGCAGCCAGCCGCGCTGGCCGCGCGAGCGGC 875
QY      424 ----- 424
Db      876 AGCCGCCGCCAGATGGCCACTTTACAAAGCTCCCGCGCTGCTGCTCCACTGCCCTT 935
QY      425 -----GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSe 441
Db      936 CCCCAGGCTCTCAAACTTATGAGCGCTGGCGGCTGGCGCGCTGGCTG 995
QY      441 rSerSerPhePheAlaAlaProAsnProTyrTyrPheSerPro----- 455
Db      996 CTCT-----TACTGGCTGTGTCACACGCCAGCCGCGCTGC 1031
QY      456 -----ThrGlyLysIleTyrProAsnThrArgLeuPro----- 466
Db      1032 CGCCGCCACCGCTGCTTACCCACCCCGCGCTTGCACGCCCTCCGCGGCTTTGG 1091
QY      467 -AlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db      1092 CGCGGTGGCGCGCTTCGACCTTGGGGGTGATTTAT 1128

RESULT 11
US-09-920-300A-1716
; Sequence 1716, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920.300A
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1716
; LENGTH: 2188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1716

```

```

Pred. No.: 6,48e-38 Length: 2188
Score: 444.00 Matches: 138
Percent Similarity: 38.36% Conservative: 45
Best Local Similarity: 28.93% Mismatches: 134
Query Match: 17.16% Indels: 161
Gaps: 16
US-09-902-772-4 (1-478) x US-09-920-300A-1716 (1-2188)

QY      122 GluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArgGln 141
Db      491 CACGCGCGCTGGGCAATTCAGAAAGAACCCCTGCTGTGGAGTGCAGCAACAGATATGCCAG 550
QY      142 TrpLeuGluThrPheValAlaLysGlyTyrGlyLeuProAspPheAlaLysLeuPheGln 161
Db      551 TGCTTCTCTGGGCGCCAAATGAGTTAGTGTGGTGAATGATGATGAGAGGTTT- 607
QY      162 AsnIleAspLysGlyLeuGlyMetThrLysAspAspPheGlnArgLeuThrPro 181
Db      608 GGCATGATGGCCAGATGCTGTATACCTTGGCAGAGACGCTTTCGAGCTGGCACT 667
QY      182 SerTrpAsnAlaAspIleLeuLeuSerHisLeu----- 192
Db      668 GACTTGTGGGTGACATTTCTGGGAGACATCTGGAGCAAAATGATCAAGAAACCAAGAA 727
QY      193 -----HisTyrLeuArg 196
Db      728 AAGACAGAGATCAATATGAAGAAATTCACACTCAGCTCCGTTCTCATTTGGATTAC 787
QY      197 GluThr----- 198
Db      788 AGCAATATCATTAGTTTGGCAGACAGAGCGCCCTATGGAATGCAGACAGAAATTAC 847
QY      199 -----ProLeu-ProHisLeuThrSerAspAspVa 208
Db      848 CCCAAAGCGGCTCTGAGACAGATGTGTCGGCCTCCACACCCAGCTACTGACTCT 907
QY      208 LasPlyAlaLeuGlnAsnSerProArgLeu----- 218
Db      908 GAGAGGAGTTTCAAGATTTCCCAAGTCTGGCTCAGCTCCGCTCAGCTCAGCTACTGC 967
QY      219 -MetHisAlaArgAsnThrGlyAlaThrPheIle-----Phe 231
Db      968 TCtGTCACTGAGGACTTCCAGAGGAGCACTTGAAATTTGCTCAACCAACTTCGGGACT 1027
QY      231 eProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgProAspLeuPr 251
Db      1028 CCCAAAGACACGACT--CCCTGAGAGACGGTGGGAGAGCTTCGAGAGCTCAGACTCC 1084
QY      251 oTyrGluGln-----AlaArgArgSerAlaTrp----- 260
Db      1085 CTCTCCAGTCTGGAGACACCGAGTCTCTGCTGTGGATGAGACAGCGGTCTTCCTTC 1144
QY      261 -----ThrSerHisSerHisProThrGlnSe 269
Db      1145 GAGAGCTTCGAAGATGACTGACGACAGCTCTCTGCTCAATAAGCAACAGATGTCTTC 1204
QY      269 rLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgProGluLe 289
Db      1205 AAGGATTACA-----TCCAAGAGAGAGAGGAGCCAGCTG-GAGCAAGGCAAAACAGTTAT 1257
QY      289 uAspProTyrGlnIleLeuGlyProThrThrSerSerArgLeuAlaAsnProGlySerGly 309
Db      1258 ACCTGACGCTGTGCTGGCGGCTTCA-----GGAAGTGAACC 1296
QY      309 nIleGlnLeuThrPglPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsnGly 329
Db      1297 TATTACAGCTGGCAGATTCTCTCGAGAGCTGTATCAGAAATTCCTGCCAGTCATTCAT 1356
QY      329 eThrTrpGluGlyThrAsnGlyLysPheLysMetThrAspProAspGluValAlaArgAr 349
Db      1357 CAGCTGAGTGTGAGAGAGATGAGATTTAAGTCTGCCGAGCCCGAGTATGAGTGGCCGCGC 1416

```

Alignment Scores:



Db 1595 -GGCGTCAGCCCGACAGCAGAGAGCTGAGTCGCCG-----GACCACCTGAGCC 1644  
 Qy 436 aleuProvalThrSerSerPhePheAlaProAsnProTyrTrpAsnSerProTh 456  
 Db 1645 GGGCCAGGCTGCTGAGCTGAGTGGGAAGCCCATCTGACACAGCTGCTCCAGAGACCCAG 1704  
 Qy 456 rGlyGly-----ThArgLeuProAlaAlaHisMetProSerHis 473  
 Db 1705 GAAGGCGAGATGTGAAATGTCCAGGAAGTGGCCAGAGACAGTGGCTTATGTGATCC 1764  
 Qy 461 oasn-----ThArgLeuProAlaAlaHisMetProSerHis 473  
 Db 1765 CAACACAGCCTCTTGACACAGCTGCTCCTGTGGCAGCAGCGCAC 1813  
 RESULT 13  
 US-09-920-300A-1693  
 ; Sequence 1693, Application US/09920300A  
 ; Patent No. US20020136728A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: King, Gordon E.  
 ; APPLICANT: Meagher, Madeleine Joy  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Secrist, Heather  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
 ; FILE REFERENCE: 210121.547  
 ; CURRENT APPLICATION NUMBER: US/09/920,300A  
 ; CURRENT FILING DATE: 2001-07-31  
 ; NUMBER OF SEQ ID NOS: 1789  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1693  
 ; LENGTH: 2268  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-920-300A-1693  
 Alignment Scores:  
 Pred. No.: 6,83e-38 Length: 2268  
 Score: 444.00 Matches: 138  
 Percent Similarity: 38.36% Conservative: 45  
 Best Local Similarity: 28.93% Mismatches: 134  
 Query Match: 17.16% Indels: 161  
 Gaps: 16  
 US-09-902-772-4 (1-478) x US-09-920-300A-1693 (1-2268)  
 Qy 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141  
 Db 571 CAGCGGGCGCTGGGCGATCCAAAGAACCCCTGCTGTGAGTGAAGCAACAGGTATGCCAG 630  
 Qy 142 TrpLeuGluTrpAlaValIleValGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161  
 Db 631 TGGCTCTCTGGGCGACCATGAGTCACTGTGTGAGCACTGAAATGTGCAGAGGCTC--- 687  
 Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181  
 Db 688 GCGATGATGCGCAGATGCTGTGTAACTTGGCAAGAACGCTTTCGAGCTTGGACCT 747  
 Qy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeu----- 192  
 Db 748 GACTTGTGGTGTGACTTCTCTGGAAACATCTGGAGCAATGATCAAGAAACCAAGAA 807  
 Qy 193 -----HisTyrLeuArg 196  
 Db 808 AAGACAGAGATCAATATGAAGAAATTCACACCTCACCTCGCTTCTCATTTGATTAAC 867  
 Qy 197 GluThr----- 198  
 Db 868 AGCATATCATATTAGTTTGGCAGAGAGAGCGCCATATGGAATGACAGACAGAAATTAC 927  
 Qy 199 -----ProLeu-ProHisLeuThrSerAspAspVa 208  
 Db 928 CCCAAGAGCGGCTCTGGACAGCATGTGTCCGGCTTCACACCCAGCGTACTCAGCTCT 987

Qy 208 lAspLysAlaLeuGlnAsnSerProArgLeu----- 218  
 Db 988 GAGACAGAGTTTCAGATTTTCCCAAGCTGTGGCTCCGCTCAGGCTACCTACTGCG 1047  
 Qy 219 -MetHisAlaArgAsnThrGlyAlaThrPheIle-----Ph 231  
 Db 1048 TCTGTCACTGAGACTTCCAGGAGCAACTTGATTTGTCTACCAACAAATTTGGGACT 1107  
 Qy 231 eProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgProAspLeuPr 251  
 Db 1108 CCCAAGACCCAGCACT---CCCTGAGAACGGTGGCGACACTTCGAGAGCTCAGACTCC 1164  
 Qy 251 oTyrGluGln-----AlaArgSerAlaTrp----- 260  
 Db 1165 CTCTCCAGCTCTGAGACAGCAGTGTCTTGTGCTGATGTGCAACAGGCTTCTCTTC 1224  
 Qy 261 -----ThrSerHisSerProThrGlnSe 269  
 Db 1225 GAGAGCTTCGAAGATGATCGACGACGCTCTCTGCTCATTAAGCCAAACATGCTTTC 1284  
 Qy 269 rLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgProGlnLe 289  
 Db 1285 AAGGATTACA-----TCCAAGAGAGAGTACCACAGTG--GAGCAAGCAACACAGTTAT 1337  
 Qy 289 uAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyG 309  
 Db 1338 ACCTGCAGCTGCTGCTGGCGGCTTCACA-----GGNAGTGGACC 1376  
 Qy 309 nIleGlnLeuTrpGlnPheLeuLeuGluLeuLeuSerAspSerAsnSerAsnGly 329  
 Db 1377 TATTCAGCTGTGGAGTTTCTCTGAGACTGCTCATACAGAAATCTCCAGTCTTCAT 1436  
 Qy 329 eThrTrpGluGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAlaArgAr 349  
 Db 1437 CAGCTGAGCTGAGACAGGATGGAGTTTAACCTGCCAGCCCGCATGGAGTGGCCGCGG 1496  
 Qy 349 gTrpGluGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuAr 369  
 Db 1497 GTGGGAGAGAGGAAATTAAGCCCAACATTAACCTACAGAGACTGACCGGGGCTTACG 1556  
 Qy 369 gTyrTyrTyrAspLysAsnIleMetThrLysValHisProProGluSerSerMetTyrLy 389  
 Db 1557 CTACTATTATGACAAAGAACATCATCCACAAAGACG----- 1590  
 Qy 389 sTyrProSerAspLeuProTyrMetSerSerTyrHisGlyLysArgTyrAlaTyrLysPh 409  
 Db 1591 -----TCGGGGAAGCGCTACGTACCGGCTT 1616  
 Qy 409 e-----AspPheHisGlyIleAlaG 416  
 Db 1617 CGTGTGGACCTCCAGAACTTGTGGGGTTTACGCCCGGAGAACTGCGACGCCATCTCG-- 1674  
 Qy 416 nAlaLeuGlnProHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAl 436  
 Db 1675 -GGCGTCCAGCCGACAGCAGGAGCTGAGTGGCGCG-----GACCACCTGAGCC 1724  
 Qy 436 aleuProvalThrSerSerPhePheAlaProAsnProTyrTrpAsnSerProTh 456  
 Db 1725 GGGCCAGGCTGCGTGTGAGTGTGGGAAGCCCATCTGACACAGCTGCTCCGAGGACCCAG 1784  
 Qy 456 rGlyGly-----IleTyrPr 461  
 Db 1785 GAAGGCGAGATTTGAAATGTCCAGGAAGTGGCCAGAGACAGTGGCTTATTTGATTC 1844  
 Qy 461 oasn-----ThArgLeuProAlaAlaHisMetProSerHis 473  
 Db 1845 CAACACAGCCTCTTGACACAGCTGCTCCTGTGTGGCAGACAGCGCAC 1893  
 RESULT 14  
 US-10-033-528-1693  
 ; Sequence 1693, Application US/10033528  
 ; Publication No. US20020131971A1





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:20:09 ; Search time 3158.2 Seconds

(without alignments) 3805.372 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKKALSVSEDSLE.....IYPNTRLPAMPHSLGTRY 478

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 24791104 segs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09902772/runat\_23072003\_093658\_14926/app\_query.fasta.1.1294  
-DB=Pending\_Patents\_NA\_Main\_QFMT-fastap -SUFFIX=crpm -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09902772.ecgn.1.1.4066.ernat.23072003\_093658\_14926 -NCPD=6 -ICPD=3  
-MMAP -LARGEDUERY -NEG.SCORES=0 -WAIT -DSPELOCK=100 -XGAPOP=10 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents\_NA\_Main:\*

1:	/cgn2_6/ptodata/1/pna/US06_PCTUS.COMB.seq:*
2:	/cgn2_6/ptodata/1/pna/US06.COMB.seq:*
3:	/cgn2_6/ptodata/1/pna/US07.COMB.seq:*
4:	/cgn2_6/ptodata/1/pna/US08.COMB.seq:*
5:	/cgn2_6/ptodata/1/pna/US081.COMB.seq:*
6:	/cgn2_6/ptodata/1/pna/US082.COMB.seq:*
7:	/cgn2_6/ptodata/1/pna/US083.COMB.seq:*
8:	/cgn2_6/ptodata/1/pna/US084.COMB.seq:*
9:	/cgn2_6/ptodata/1/pna/US085.COMB.seq:*
10:	/cgn2_6/ptodata/1/pna/US086.COMB.seq:*
11:	/cgn2_6/ptodata/1/pna/US087.COMB.seq:*
12:	/cgn2_6/ptodata/1/pna/US088.COMB.seq:*
13:	/cgn2_6/ptodata/1/pna/US089.COMB.seq:*
14:	/cgn2_6/ptodata/1/pna/US090.COMB.seq:*
15:	/cgn2_6/ptodata/1/pna/US091.COMB.seq:*
16:	/cgn2_6/ptodata/1/pna/US092.COMB.seq:*
17:	/cgn2_6/ptodata/1/pna/US093.COMB.seq:*
18:	/cgn2_6/ptodata/1/pna/US094.COMB.seq:*
19:	/cgn2_6/ptodata/1/pna/US095.COMB.seq:*
20:	/cgn2_6/ptodata/1/pna/US095B.COMB.seq:*
21:	/cgn2_6/ptodata/1/pna/US095C.COMB.seq:*
22:	/cgn2_6/ptodata/1/pna/US095D.COMB.seq:*
23:	/cgn2_6/ptodata/1/pna/US096.COMB.seq:*
24:	/cgn2_6/ptodata/1/pna/US096B.COMB.seq:*
25:	/cgn2_6/ptodata/1/pna/US096C.COMB.seq:*
26:	/cgn2_6/ptodata/1/pna/US096D.COMB.seq:*
27:	/cgn2_6/ptodata/1/pna/US096E.COMB.seq:*
28:	/cgn2_6/ptodata/1/pna/US097A.COMB.seq:*

29:	/cgn2_6/ptodata/1/pna/US097B.COMB.seq:*
30:	/cgn2_6/ptodata/1/pna/US097C.COMB.seq:*
31:	/cgn2_6/ptodata/1/pna/US098A.COMB.seq:*
32:	/cgn2_6/ptodata/1/pna/US098B.COMB.seq:*
33:	/cgn2_6/ptodata/1/pna/US098C.COMB.seq:*
34:	/cgn2_6/ptodata/1/pna/US099A.COMB.seq:*
35:	/cgn2_6/ptodata/1/pna/US099B.COMB.seq:*
36:	/cgn2_6/ptodata/1/pna/US099C.COMB.seq:*
37:	/cgn2_6/ptodata/1/pna/US099D.COMB.seq:*
38:	/cgn2_6/ptodata/1/pna/US100A.COMB.seq:*
39:	/cgn2_6/ptodata/1/pna/US100B.COMB.seq:*
40:	/cgn2_6/ptodata/1/pna/US101A.COMB.seq:*
41:	/cgn2_6/ptodata/1/pna/US101B.COMB.seq:*
42:	/cgn2_6/ptodata/1/pna/US102A.COMB.seq:*
43:	/cgn2_6/ptodata/1/pna/US102B.COMB.seq:*
44:	/cgn2_6/ptodata/1/pna/US6000.COMB.seq:*
45:	/cgn2_6/ptodata/1/pna/US6001.COMB.seq:*
46:	/cgn2_6/ptodata/1/pna/US6002.COMB.seq:*
47:	/cgn2_6/ptodata/1/pna/US6003.COMB.seq:*
48:	/cgn2_6/ptodata/1/pna/US6004.COMB.seq:*
49:	/cgn2_6/ptodata/1/pna/US6005.COMB.seq:*
50:	/cgn2_6/ptodata/1/pna/US6006.COMB.seq:*
51:	/cgn2_6/ptodata/1/pna/US6007.COMB.seq:*
52:	/cgn2_6/ptodata/1/pna/US6008.COMB.seq:*
53:	/cgn2_6/ptodata/1/pna/US6009.COMB.seq:*
54:	/cgn2_6/ptodata/1/pna/US6010.COMB.seq:*
55:	/cgn2_6/ptodata/1/pna/US6011.COMB.seq:*
56:	/cgn2_6/ptodata/1/pna/US6012.COMB.seq:*
57:	/cgn2_6/ptodata/1/pna/US6013.COMB.seq:*
58:	/cgn2_6/ptodata/1/pna/US6014.COMB.seq:*
59:	/cgn2_6/ptodata/1/pna/US6015.COMB.seq:*
60:	/cgn2_6/ptodata/1/pna/US6016.COMB.seq:*
61:	/cgn2_6/ptodata/1/pna/US6017.COMB.seq:*
62:	/cgn2_6/ptodata/1/pna/US6018.COMB.seq:*
63:	/cgn2_6/ptodata/1/pna/US6019.COMB.seq:*
64:	/cgn2_6/ptodata/1/pna/US6020.COMB.seq:*
65:	/cgn2_6/ptodata/1/pna/US6021.COMB.seq:*
66:	/cgn2_6/ptodata/1/pna/US6022.COMB.seq:*
67:	/cgn2_6/ptodata/1/pna/US6023.COMB.seq:*
68:	/cgn2_6/ptodata/1/pna/US6024.COMB.seq:*
69:	/cgn2_6/ptodata/1/pna/US6025.COMB.seq:*
70:	/cgn2_6/ptodata/1/pna/US6026.COMB.seq:*
71:	/cgn2_6/ptodata/1/pna/US6027.COMB.seq:*
72:	/cgn2_6/ptodata/1/pna/US6028.COMB.seq:*
73:	/cgn2_6/ptodata/1/pna/US6029.COMB.seq:*
74:	/cgn2_6/ptodata/1/pna/US6030.COMB.seq:*
75:	/cgn2_6/ptodata/1/pna/US6031.COMB.seq:*
76:	/cgn2_6/ptodata/1/pna/US6032.COMB.seq:*
77:	/cgn2_6/ptodata/1/pna/US6033.COMB.seq:*
78:	/cgn2_6/ptodata/1/pna/US6034.COMB.seq:*
79:	/cgn2_6/ptodata/1/pna/US6035.COMB.seq:*
80:	/cgn2_6/ptodata/1/pna/US6036.COMB.seq:*
81:	/cgn2_6/ptodata/1/pna/US6037.COMB.seq:*
82:	/cgn2_6/ptodata/1/pna/US6038.COMB.seq:*
83:	/cgn2_6/ptodata/1/pna/US6039.COMB.seq:*
84:	/cgn2_6/ptodata/1/pna/US6040.COMB.seq:*
85:	/cgn2_6/ptodata/1/pna/US6041.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2436	94.1	1528	34	US-09-902-772-3
2	2262.5	87.4	1447	34	US-09-902-772-1
3	2195.5	84.8	3309	25	US-09-652-109-9617
4	2195.5	84.8	3309	28	US-09-716-953-1936
5	2195.5	84.8	3309	29	US-09-721-589-5498
6	2195.5	84.8	3309	29	US-09-726-806-4625

```

7 2195.5 84.8 3309 29 US-09-726-811-4920 Sequence 4920, Ap
8 2180.5 84.3 1389 35 US-09-948-941-233 Sequence 233, App
9 2180.5 84.3 3166 1 PCT-US02-23913-98 Sequence 98, Appl
10 2180.5 84.3 3166 30 US-09-784-356-26 Sequence 26, Appl
11 2180.5 84.3 3166 38 US-10-021-660-26 Sequence 26, Appl
12 2180.5 84.3 3166 42 US-10-205-823-98 Sequence 26, Appl
13 2017.5 78.0 2172 80 US-09-360-207-355 Sequence 355, App
14 1808.5 69.9 3126 18 US-09-442-384A-475 Sequence 475, App
15 1808.5 69.9 3126 18 US-09-442-384B-475 Sequence 475, App
16 1482.5 57.3 1359 38 US-10-035-832-1210 Sequence 1210, App
17 1482.5 57.3 1359 7 US-10-052-482-198 Sequence 198, App
18 1482.5 57.3 2938 7 US-08-343-443-3 Sequence 3, Appl1
19 1482.5 57.3 2937 1 PCT-US02-18947-710 Sequence 710, App
20 1482.5 57.3 2957 38 US-10-007-926A-295 Sequence 295, App
21 1482.5 57.3 2957 38 US-10-035-832-1209 Sequence 1209, App
22 1482.5 57.3 2957 39 US-10-052-482-197 Sequence 197, App
23 1482.5 57.3 2957 41 US-10-172-118-710 Sequence 710, App
24 1482.5 57.3 3298 25 US-09-652-814-9467 Sequence 9467, App
25 1482.5 57.3 3298 27 US-09-699-899-6013 Sequence 6013, App
26 1482.5 57.3 3298 28 US-09-716-953-1954 Sequence 1954, App
27 1482.5 57.3 3298 28 US-09-721-589-5554 Sequence 1139, App
28 1482.5 57.3 3298 29 US-09-721-589-5554 Sequence 5554, App
29 1482.5 57.3 3298 29 US-09-726-171-2148 Sequence 2148, App
30 1482.5 57.3 3298 29 US-09-726-806-4681 Sequence 4681, App
31 1482.5 57.3 3372 17 US-09-396-870-7615 Sequence 7615, App
32 1474.5 57.0 1359 38 US-10-035-832-1207 Sequence 1207, App
33 1474.5 57.0 1359 38 US-10-052-482-195 Sequence 195, App
34 1474.5 57.0 1729 14 US-09-053-375B-627 Sequence 627, App
35 1474.5 57.0 1729 38 US-10-035-832-1206 Sequence 1206, App
36 1474.5 57.0 1729 39 US-10-052-482-194 Sequence 194, App
37 1474.5 57.0 1729 80 US-09-360-207-10845 Sequence 10845, App
38 1469.5 56.8 3110 76 US-09-324-185-33223 Sequence 33223, App
39 1468.5 56.7 1307 61 US-09-172-373-5267 Sequence 5267, App
40 1394.5 53.9 1932 1 PCT-US02-25766-2596 Sequence 2596, App
41 1394.5 53.9 1932 14 US-09-053-375B-327 Sequence 327, App
42 1394.5 53.9 1932 18 US-09-442-384A-530 Sequence 530, App
43 1394.5 53.9 1932 18 US-09-442-384B-530 Sequence 530, App
44 1061.0 41.0 661 29 US-09-726-806-3734 Sequence 3434, App
45 1019.0 39.4 840 25 US-09-652-109-7574 Sequence 7574, App

```

## ALIGNMENTS

```

RESULT 1
US-09-902-772-3
: Sequence 3, Application US/09902772
: GENERAL INFORMATION:
: APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
: FILE REFERENCE: chugai selyaku kabushiki kaisha 5001
: CURRENT FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: US/09/902,772
: PRIOR FILING DATE: 1997-06-18
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 1528
: TYPE: DNA
: ORGANISM: C-ery gene, chicken DNA
US-09-902-772-3

```

```

Alignment Scores:
Pred. No.: 4,41e-178 Length: 1528
Score: 2436.00 Matches: 458
Percent Similarity: 91.97% Conserves: 0
Best Local Similarity: 91.97% Mismatches: 0
Query Match: 94.13% Indels: 40
DB: 34 Gaps: 2
US-09-902-772-4 (1-478) x US-09-902-772-3 (1-1528)

```

```

Qy 1 MetalaserthrilelysgluAlaIleuSerValValSerluaspInSerLeupehlu 20
Db 63 ATGCAGACACTATTAAAGAGCATTTATCACTGTGATGTAAGACCAAGCTCTTGTAG 122
Qy 21 CysAlaIyrglySerProHisIleuAlaIythrGluMetThrAlaSerSerSerGlu 40
Db 123 TGTCCCTACGAGATGCCGCCACCTTGCAAGACAAATGACAGGCTCTCTTCCAGTAA 182
Qy 41 TyrcIyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGluPro 60
Db 183 TATGGCAACATCAATMAAGATGAGCCCGCGCTTCCACGACGAGCTGTATTCACACCC 242
Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db 243 CCGGCCAGAGTTACCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
Qy 81 SerProaspAspCysSerSerValAlaIyGlyGlyIyMetValSerSerSerAspAsnVal 100
Db 303 TCACCTGATGCTACGCTGACGCTGCAAAAGAGGAAATGTTACGATTCAGACAAATGTT 362
Qy 101 GlyMetAsnTyrcIySerTyrcMetGluGluIyHisIleProProAsnMetThrThr 120
Db 363 GGGATGAACTATGGAAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
Qy 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
Db 423 AATGACGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
Qy 141 GlnTrpLeuGlnTrpAlaValIyGluIyGluIyGluIyGluIyGluIyGluIyGluIyGlu 160
Db 483 CAGTGGCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
Qy 161 GlnAsnIleAspGlyIyGluIyGluIyGluIyGluIyGluIyGluIyGluIyGluIyGlu 180
Db 543 CAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
Qy 181 ProSerTyrcAsnAlaAspIleLeuLeuSerHisIleuHisIyTrpLeuArgGluThrPro 200
Db 603 CCGAGCTATACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
Qy 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
Db 663 CCACATTTGACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
Qy 221 AlaArgAsnThrGlyIyAlaIyThrPheIlePheProAsnThrSerValTyrcProGluAla 240
Db 723 GCTAGAAACACAGAGAGGACGACCTTTATTTTCCAAATACATCAGTTTACCAGAGCA 782
Qy 241 ThrGlnArgIleThrThrArgProAspLeuProTyrcIyGlnAlaIyArgSerAlaTrp 260
Db 783 ACGCAAGAAATACACACAGGCGCAGATTTACCTTATGACGAGAGGAGAGATCAGCTGG 842
Qy 261 ThrSerHisSerHisProThrGlnSerLysAlaIyThrGlnProSerSerThrValPro 280
Db 843 ACGAGTACAGCCATCCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCC 902
Qy 281 LysThrGlnLyspGlnArgProGlnIleuAspProTyrcIyGlnIleuGlnIyProThrSer 300
Db 903 AAAACAGACAGACGCGCTCCAGTTAGATTCCTTATCAGATTTTGGACCGACAGCAGC 962
Qy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheIleuLeuLeuLeu 320
Db 963 CGTCTTCCAAATCCAGGAGTGGGACATACATACATACATACATACATACATACATACAT 1022
Qy 321 SerAspSerSerAsnSerAsnGlyIleThrTrpGluIyGlnIyGlnIyGlnIyGlnIyGln 340
Db 1023 TCGGACAGCTCCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 1082
Qy 341 ThrAspProaspGluValAlaIyArgArgTrpGlyIyGlnIyGlnIyGlnIyGlnIyGln 360
Db 1083 ACAGACCCCTATGAAAGTGGCTCGCGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1142
Qy 361 TyrAspLysLeuSerArgAlaLeuArgTyrcTyrcTyrcAspLysAsnIleMetThrLysVal 380

```



```

|||||
Db      1143 TATGCAAACTGAGCGGTGACTGTACTAGTACAAATAATATATGCTAACTT 1202
QY      380 -----
Db      1203 CATGGTAAACGTATGCTTACAAATTTGATTTCCACGGAAWCGCTCAGGCCCTCCAGCCT 1262
QY      381 HisProgluserSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db      1263 CACCTCCAGATCATCTCCATGACAAATACCATGACCTCCCTACATAGATTCCTAC 1322
QY      401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1322 -----
QY      421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db      1323 CATGCACACCCCAAGAAATGAACTTTGAGCTCCCATCCCGCTTGGCCGTAC 1382
QY      441 SerSerSerPhePheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyLysTyr 460
Db      1383 TCATCCAGCTTTTGGCTGCCCCCTAATCCATCTGGAATTCACCAACTGGAGGCAATCTAC 1442
QY      461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db      1443 CCCAATACGAGCTGCCAGCTGCTCATATGCTTCCCTTGGCAGCCTACTAC 1496

```

## RESULT 2

```

US-09-902-772-1
; Sequence 1, Application US/09902772
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; FILE REFERENCE: the proteins
; CURRENT APPLICATION NUMBER: US/09/902,772
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US/08/878,177
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1447
; TYPE: DNA
; ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA
US-09-902-772-1

```

## Alignment Scores:

```

Pred. No.: 1,07e-164      Length: 1447
Score: 2262.50           Matches: 430
Percent Similarity: 86.35%      Conservative: 0
Best Local Similarity: 86.35%      Mismatches: 1
Query Match: 87.42%           Indels: 67
DB: 34                    Gaps: 3

```

US-09-902-772-4 (1-478) x US-09-902-772-1 (1-1447)

```

QY      1 MetAlaSerThrIleLysGlnAlaLeuSerValAlaSerGluAspGlnSerLeuPheGln 20
Db      63 ATGGCAAGCACTAATTAAAGAGACATTTATCACTGGTGAAGAACACACTGCTTTGAG 122
QY      21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40
Db      123 TGTGCTAGCGATGCGCCCACTTGCAAGACAGAAATGACAGCCCTCTTCCAGTAA 182
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
Db      183 TATGGCAAAACATCAAAAGATGAGCCGCGCTTCCCGACGAGACTGTTATTCACAGGCC 242
QY      61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db      243 CCGGCCAGAGTTACCATTAAGATGAGTGTAAACCAACCAAGTAAATGCGTCAAGGAAAT 302

```

```

QY      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db      303 TCACCTGATGACTGACGCGCTGGCAGAAAGAGGAAATGCTTACGACGTTCCAGACATGTT 362
QY      101 GlyMetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrThr 120
Db      363 GGGATGACATATGGAAGCTACATGGAAGAAGACATATCCGCCCTCCAAATATGACAAACC 422
QY      121 AsnGluArgValIleValProAlaAspProThrLeuTyrPserThrAspHisValArg 140
Db      423 AATGAAAGCAAGATTAATGTCACAGCAATCTCCTGATATGAGCAGACCATGTAGCG 482
QY      141 GlnThrLeuGlnUTRPalValLysGlnLysTyrLeuProAspValAspIleLeuLeuPhe 160
Db      483 CAGTGGCTGGAGTGGGCGAGTGAAGATGATGCTTCCAGACGTGACATCTGTTGTTTC 542
QY      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db      543 CAGAACATTTGATGGAAAGAGTTGTGTAATGACCAAGATGACTTCACAGACTCAGC 602
QY      181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
Db      603 CCGAGCTTAACGCGAGATATCTCTCTGACACCTACCTACCTCAGAGAGA----- 656
QY      201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
Db      656 -----
QY      221 AlaArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGlnAla 240
Db      657 -----GGAGCCACTTTATTTTCCAAATACATACAGTTTACCAGAAACA 701
QY      241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgSerAlaTyr 260
Db      702 ACGCAAAAGATTAACAAACAGCCAGATTTACCTTATGAGCAAGCAGAGAGATTCAGCGTGG 761
QY      261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db      762 ACGAGTCAACAGCCATCCACCTGAGTCAAAACCTCCCAACATCATCTTCAACAGTGGCC 821
QY      281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db      822 AAACAGAGAGACAGCGCTCCAGTATATCTTATGATTCCTTGGACGACGACGACG 881
QY      301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuLeuGlnLeuLeu 320
Db      882 CGTCTTGCAAATCCAGGAGTGGGACAGATACAGCTATGAGCAGATTCCTACTGAGCTTCTG 941
QY      321 SerAspSerSerAsnSerAsnCysIleThrThrProGlnGlyThrAsnGlyLysPheLysMet 340
Db      942 TCGGACAGCTCACTCCACACTGCATGCATCCCTGGGAGGAGGACAAATGGGAGTTCAAGATG 1001
QY      341 ThrAspProAspGlnValAlaArgArgTyrProGlyLysSerLysProAsnMetAsn 360
Db      1002 ACAGACCTGATGAAGTGGCTCGGGCTGGGAGAGAGAGAAACCAACCTAACATGAGAC 1061
QY      361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db      1062 TATGCAAAACTCAGCCGCTGCTCTGCTACTACTATGACAAATAATATATGACTAAAGTT 1121
QY      380 -----
Db      1122 CATGGTAAACGTATGCTTACAAATTTGATTTCCACGGAATCGCTCAGGCCCTCCAGCCT 1181
QY      381 HisProProgluserSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db      1182 CACCTCCAGATCATCTCCATGACAAATACCATGACACCTCCCTCATATGATTCCTAC 1241
QY      401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1241 -----
QY      421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440

```

Db 1242 CATCACACCCCCGAGAGTAACTTTGTACCTCCCACTCCCTGCTTTCCCGTAACC 1301  
 Oy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyLeuTyr 460  
 Db 1302 TCATCCAGCTTTTGTGCTGCCCCCTAATCCATCTGGAATTCACCAACTGGAGGACTTAC 1361  
 Oy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478  
 Db 1362 CCCAATACCAAGCTGCCAGCTGCTCATATGCTTCCCATCTTGGACCTACTAC 1415

## RESULT 3

US-09-652-109-9617  
 ; Sequence 9617, Application US/09652109  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McCarthy, Sean A.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
 ; TITLE OF INVENTION: THEREFOR  
 ; FILE REFERENCE: 1600,1180-001  
 ; CURRENT APPLICATION NUMBER: US/09/652,109  
 ; PRIOR FILING DATE: 2000-08-30  
 ; PRIOR APPLICATION NUMBER: 60/151,128  
 ; PRIOR FILING DATE: 1999-08-30  
 ; NUMBER OF SEQ. ID NOS: 10105  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9617  
 ; LENGTH: 3309  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-652-109-9617

## Alignment Scores:

Pred. No.: 4,4e-159 Length: 3309  
 Score: 2195.50 Matches: 416  
 Percent Similarity: 84.97% Conservative: 8  
 Best Local Similarity: 83.37% Mismatches: 10  
 Query Match: 84.83% Indels: 65  
 Db: 25 Gaps: 4

US-09-902-772-4 (1-478) x US-09-652-109-9617 (1-3309)

Oy 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 128 ATGGCCAGCACTATTAAAGAGCCTTATTCAGTTGTGATGGAGCAAGTCTTGTTCAG 187  
 Oy 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 Db 188 TGTGCTACGAGCAAGCCACACCTGGTAAGACAGATGACCGGCTCTCTCCAGCGAC 247  
 Oy 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 Db 248 TATGACAGACTTCCAGATGAGCCACGCTCCCTCAGCAGATGGCTGTCTCAACCC 307  
 Oy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 308 CCACGCCAGGCTCACATCAAAATGTAACCTTCCCGAGGTGAATGGCTCAAGGAC 367  
 Oy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 Db 368 TCTCTCATGATGATCAGTGTGGCCAAAGCGGAGAAAGTGTGGGACGCCAGACACGTT 427  
 Oy 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120  
 Db 428 GGGATGAACCTACGCGACCTACATGAGAGAGACACATGACACCCCAACATGACACAG 487  
 Oy 121 AsnGlnArgArgValIleValAlaProAlaAspProThrIleuTrpSerThrAspHisValArg 140  
 Db 488 AACGAGGCGAGATTATCTGCGCAGAGATCTTACGCTATGAGATAGACACATGCTCGG 547  
 Oy 141 GlnTrpLeuGlnTrpAlaValLysGlyTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 Db 548 CAGTGGCTGAGTGGGCGGCTCAAGAAATATGGCTTCCACAGCTCAACATCTTATATTC 607

Oy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 608 CAGAAACATCATGATGGAGGAAGACTGTGCAAGATGACCAAGAGCAGCTTCCAGAGCTCAC 667  
 Oy 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200  
 Db 668 CCCAGCTAACAGCGCGCATCTCTCTCATATCTCCACATCCACAGACACCTCTCTT 727  
 Oy 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220  
 Db 728 CCACATTTGACTTATGATGATGATGATGATTAACCTTACAAACCTTCCACGGTTATATCAT 787  
 Oy 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240  
 Db 788 GCTAGAAAACACA----- 799  
 Oy 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrp 260  
 Db 800 -----GATTTACCATATGAGCCCCCAGAGAGATGACCTGG 835  
 Oy 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThrVal 279  
 Db 836 ACCGGTACAGGCGCACCCACAGCCCGCAGTCCGAAAGCTGTCAACCATCTCTCCACAGTG 895  
 Oy 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299  
 Db 896 CCCAAACTGAAGACACGAGCTCTCAGTATGATTCATGATTCATGATTCGACCAACAGT 955  
 Oy 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeu 319  
 Db 956 AGCCGCTTGCAAAATCCAGGAGTGGCCAGATCCAGCTTGGGAGTTCTCTCTGAGCTC 1015  
 Oy 320 LeuSerAspSerSerAsnSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLys 339  
 Db 1016 CTGTGGAGACAGCTCAACTCAGCTGATCATCTGGAAGGACCAAGGAGGAGTTCAAG 1075  
 Oy 340 MetThrAspProAspGluValAlaArgArgTrpGlyLysLysSerLysProAsnMet 359  
 Db 1076 ATGACGATCCCGACGAGTGGCGCGCGCTGGGAGAGCGGAGACCAACCCCAACATG 1135  
 Oy 360 AspTyrAspLysLeuSerArgAlaLeuArgTyrTyrTrpAspLysAsnIleMetThrLys 379  
 Db 1136 AACTACGATTAAGCTCAGCGCGGCTCCGTTACTATGATGACAAAGATCATATGACCAAG 1195  
 Oy 380 ValHisProGlnIleSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399  
 Db 1196 GTC----- 1198  
 Oy 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419  
 Db 1199 ---CATGGGAAGCGCTACGGCTACAGTTGCACTTCCACGAGGATGCCAGGCCCTCAG 1255  
 Oy 420 Pro----- 420  
 Db 1256 CCCACCCCCCGGAGTATCTGTACAGTACCCCTCAGACCTCCCTGATACATGGGCTCC 1315  
 Oy 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439  
 Db 1316 TATACGCCCCACCCAGAGAGATGAACTTGTGGCGCCACCCCTCAGGCTCCCGCTG 1375  
 Oy 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLys 459  
 Db 1376 ACATCTTCCAGTTTGTGCTGCCCCCAACCCATATCGGAATTCACCAACTGGGGGTATA 1435  
 Oy 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478  
 Db 1436 TACCCCAACACTAGCTCCACACGCAATATGCTTCTCATCTGCGGCTTACTAC 1492

## RESULT 4

US-09-716-953-1936  
 ; Sequence 1936, Application US/09716953  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gearling, David P.

```

; APPLICANT: Holtzman, Douglas A.
; APPLICANT: McCarthy, Sean
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2044-001
; CURRENT APPLICATION NUMBER: US/09/716,953
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/167,413
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 2620
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1936
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-716-953-1936

Alignment Scores:
Pred. No.: 4,4e-159 Length: 3309
Score: 2195.50 Matches: 416
Percent Similarity: 84.978 Conservative: 8
Best Local Similarity: 83.378 Mismatches: 10
Query Match: 84.83% Indels: 65
DB: 28 Gaps: 4

US-09-902-772-4 (1-478) x US-09-716-953-1936 (1-3309)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB 128 ATGGCCACGACATATTAAGGAAGCCCTTATCAGTTGGATGAGAGACCACTCGTTGTTGAG 187
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerL 40
DB 188 TGTGCTTACGGAGACCCACCTGGCTAAGACAGAGATGACCGCTCTCTCCAGCCAGC 247
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
DB 248 TATGACAGACCTCCCAATGAGCCACGCGCTCCCTCAGCAGATGGCTGTCTCAACCC 307
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 308 CCAGCCAGGCTGCACCATCAAAATGGAAATGAACCTTACCCAGTGAATGGCTCAAGAAC 367
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnVal 100
DB 368 TCTCTGTGATGATGAGTGTGGCCAAAGCGGAAGATGGTGGCGACCCAGACACCGTT 427
QY 101 GlyMetAsnTyrGlySerTyrMetGlnGluLysHisIleProProProAsnMetThrThr 120
DB 428 GGGATGACACTACGAGCTACATGAGAGAGACCAATGCCACCCCAACATGACCCAGC 487
QY 121 AsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
DB 488 AACGGCCAGAGTATGCTCCAGCAGATCTTACGCTATGAGATACAGACCATGTGGCG 547
QY 141 GlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuPhe 160
DB 548 CAGTGGCTGGAGTGGCGGTGAAGAATATGCGCTCCAGACGTCACATCTTGTTATTC 607
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB 608 CAGAAACATCGATGGGAAGAACTGTGCAAGATGACCAAGACGACTTCCAGAGGTCCAC 667
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB 668 CCCAGCTCAACAGCGCGCATCTCTCTCACATCTCCACTCCACAGAGACATCTCTT 727
QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
DB 728 CCACATTTGACTTCAGATGATGTGATTAAGCCTTACAAACTCTCCAGGTTAATGAT 787
QY 221 AlaArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240

```

```

DB 788 GCTAGAAACACA----- 799
QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnAlaArgSerAlaTrp 260
DB 800 -----GATTACATATGAGACCCCAAGAGATCAGCTGG 835
QY 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 279
DB 836 ACCGGTCAAGGCGCACCCAGCCAGAGTGAAGCTGCTCAACCATCTCTCCACAGATG 895
QY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
DB 896 CCCAAATCTGAAGACACCGCTCCTCAGTTAGATCTTATCAGATTTCTTGACCAACAGT 955
QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 319
DB 956 AGCCGCTTGCAAAATCCAGAGCAGTGGCAGATCCAGCTTGTGCACTCTCTCGAGACTC 1015
QY 320 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyGluPheLys 339
DB 1016 CTGTGGACACTCCACATCCAGCTGATCACTGGAAGGACCAACGGGAGTTCAAG 1075
QY 340 MetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 359
DB 1076 ATGACGATCCCGACGAGAGTGGCGCGCTGGGAGAGCGGAAGACCAACCAACATG 1135
QY 360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
DB 1136 AACTACGATTAAGCTACGCCGCGCTCGTTACTACTATGACAAAGATCATGACCAAG 1195
QY 380 ValHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399
DB 1196 GTC----- 1198
QY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
DB 1199 ---CATGGGAAGCGTTAGCGCTACAGATGTCAGCTCCACGGGATGGCCGCTCCAG 1255
QY 420 Pro----- 420
DB 1256 CCCCACCCCGGAGTCACTCTGTACAAAGTACCCCTCAGACCTCCCTACATGGGCTCC 1315
QY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProAl 439
DB 1316 TATCAGCGCCACCCACAGAAAGTGAATTTGGGGCCCAACCTCCAGCCCTCCGCTG 1375
QY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIle 459
DB 1376 ACATCTTCAGATTTTGTGGTGGCCCAAAACCATATCTGGAATTCACCAACTGGGGCTATA 1435
QY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
DB 1436 TACCCCAACACTAGGCTCCACAGCATATGCTCTCACTGTGGGCACTTACTAC 1492

RESULT 5
US-09-721-589-5498
; Sequence 5498, Application US/09721589
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Villaveal, Jean-Luc
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2045-001
; CURRENT APPLICATION NUMBER: US/09/721,589
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 60/167,380
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 7017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5498
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-09-721-589-5498

## Alignment Scores:

Pred. No.: 4,4e-159  
 Score: 2195.50  
 Percent Similarity: 84.978  
 Best Local Similarity: 83.378  
 Query Match: 84.83%

DB: 29

Length: 3309  
 Matches: 416  
 Conservative: 8  
 Mismatches: 10  
 Indels: 65  
 Gaps: 4

US-09-902-772-4 (1-478) x US-09-721-589-5498 (1-3309)

```

QY      1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      128 ATGGCCACACTATTAAAGGAAGCCCTTATCAGTTGTGATGAGGACCACTCGCTTTGAG 187
QY      21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB      188 TGTGCTTACGGAGACCCACACTGCTAGACAGATGACCGCTCTCTCCAGCGAC 247
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
DB      248 TATGACAGACTTCCAAAGATGAGCCAGCGCTCCCTCAGCAGAGATTGGCTGTCAACCC 307
QY      61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB      308 CCAGCCAGGCTCACCATCAAAATGGAATGTAACCTACCCAGGTGAATGGCTCAAGAAC 367
QY      81 SerProAspAspCysSerSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB      368 TCTCTGTATGATGACAGTGTGGCCCAAGGCGGAGATGTTGGGCAAGCCAGACACCGTT 427
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
DB      428 GGGATGACCTACGAGCAGTACATGAGAGAGAACACATGCCACCCCAACATGACCAAG 487
QY      121 AsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
DB      488 AACGGCCCAAGTATGTCGCCACAGATCTCAGCTATGAGATGACCATGTGCGG 547
QY      141 GlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB      548 CAGTGGCTGGAGTGGCGGCTCAAAACATATGGCTTCCAGACGTCAACATCTTGTATTC 607
QY      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB      608 CAGAACATCGATGGAGAGAACTGTCAAAGATGACCAAGACGACTTCCAGAGGCTCAC 667
QY      181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB      668 CCCAGCTACAAAGCCGACATCTCTCTCAATCTCCACTCCTCAGAGAGACTCTCTTT 727
QY      201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
DB      728 CCACATTTGACTTCAGATGATGTGATTAAGCTTACAAACTCTCCAGCGTTAATGCAT 787
QY      221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
DB      788 GCTAGAAACACA----- 799
QY      241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 260
DB      800 -----GATTACCATATGAGCCCCCGAGAGATCAACCCGCG 835
QY      261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 279
DB      836 ACCGGTACGGCCACCCACGCCCGCAGTCAAGTGTCTCAACCATCTCTCCACACATG 895
QY      280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
DB      896 CCCAAACTGAAGACACAGCTCTCATAGTTAGATCTTCAAGATCTTGGAGCAACAGAGT 955
QY      300 SerArgLeuAlaAsnProGlySerGlyLysIleGlnLeuTrpGlnPheLeuLeuGluLeu 319

```

```

DB      956 AGCCGCTTGGCAAAATCCAGGAGTGCCAGATCCAGCTTGGCACTTCTCTGGAGCTC 1015
QY      320 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLupheLys 339
DB      1016 CTGTGGAGACGTCCAACTCCAGCTGCATCTACCTGGGAAGGACCAAGGAGTTCAG 1075
QY      340 MetThrAspProAspGluValAlaArgArgTrpGlyLysArgLysSerLysProAsnMet 359
DB      1076 ATGACGATATCCGACGAGAGTGCGCGCGCTGGGAGAGAGCGGAAGCAACCAACATG 1135
QY      360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
DB      1136 AACTACGATTAAGCTTACGCGCGCTCCCTCACTACTATGACAAAGATCATGACCAAG 1195
QY      380 ValHisProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSer 399
DB      1196 GTC----- 1198
QY      400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
DB      1199 ---CATGGGAAGCGCTACGCTACAAAGTTGCACCTTCCAGGGATGCCAGGCCCTCAG 1255
QY      420 Pro----- 420
DB      1256 CCCACCCCCCGAGTCACTCTGTACAGATACCCCTAGACCTTCCGTACATGGGCTCC 1315
QY      421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439
DB      1316 TATCAGCGCCACCCACAGATGACTTTGTGGGCGCCCAACCTCCAGCCCTCCCGTG 1375
QY      440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIle 459
DB      1376 ACATCTTCCAGTTTGTGTGTCGCCCAAAACCATGTAATTCACAACTGGGGGTATA 1435
QY      460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTrp 478
DB      1436 TACCCCAACACTAGGCTCCCAACGACCATATGCTTCATCTGGGACCTTACTAC 1492

RESULT 6
US-09-726-806-4625
; Sequence 4625, Application US/09726806
; GENERAL INFORMATION:
; APPLICANT: Galvin, Katherine
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600, 2028-001
; CURRENT APPLICATION NUMBER: US/09/726, 806
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/168, 135
; NUMBER OF SEQ ID NOS: 6283
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4625
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-806-4625

Alignment Scores:
Pred. No.: 4,4e-159
Score: 2195.50
Percent Similarity: 84.978
Best Local Similarity: 83.378
Query Match: 84.83%

Length: 3309
Matches: 416
Conservative: 8
Mismatches: 10
Indels: 65
Gaps: 4

US-09-902-772-4 (1-478) x US-09-726-806-4625 (1-3309)
QY      1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      128 ATGGCCACACTATTAAAGGAAGCCCTTATCAGTTGTGATGAGGACCACTCGCTTTGAG 187

```

```

QY 21 CysAlaTyrGlySerProHISLeuAlaLysThrGluMetThrAlaSerSerSerglu 40
DB 188 TGTGCTACGGAAGCCACACCTGGCTTAACAGAGATGACCCGCTCTCTCCAGCGAC 247
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerglnPro 60
DB 248 TATGAGACAGACTTCCAAAGATGAGAGCCAGCGCTCCCTCAGCAGATGGCTGTCTCAACC 307
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 308 CCAAGCCAGGGTCCACCAAAATGGAATGTACCTTACCCAGCGTGAATGGCTCAAGGAC 367
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB 368 TCTCCTGATCAATGCACTGTGGCCAAAGCGGAGATGGTGGCGACCCAGACACCGCTT 427
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
DB 428 GGGATGAACTACGCGACCTACATGAGAGAGACACATGCCACCCCAACATGACACAG 487
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
DB 488 AACGAGCGCAGAGTTATCGTCCAGCAGATCTTACCGTATGAGTACAGACATGTGCGG 547
QY 141 GlnTyrLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB 548 CAGTGGCTGAGTGGGGGGGTAAGAAATATGGCTTCCAGACCTCAACATCTTGTATTC 607
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB 608 CAGAACATCGATGGGAAGAACTGTGCAGATGACCAAGAGACGACTTCCAGAGCTCACC 667
QY 181 ProSerTyrAsnAlaAspIleLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB 668 CCAAGCTACACCCGACATCTTCTCTCATCTCCATCTCCAGACGAGACTCTCTT 727
QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
DB 728 CCACATTTGACTTCAGATGATGTGTGAATAAAGCCTTACAAACCTCCACAGCTTAATGCAT 787
QY 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrServalTyrProGluAla 240
DB 788 GCTGAAGAACACA----- 799
QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgAsnAlaTyr 260
DB 800 -----GATTACCATATGAGCCGCCAGGAGATCAGCTGG 835
QY 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 279
DB 836 ACCGGTCACGGCCACCCAGCCGACGTCGAAAGCTGCTCAACCATCTCTTCCACAGTG 895
QY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
DB 896 CCCCAGAACTGAAACACAGCGCTCCTCAGTATCTTATCATAGATCTTGGACCAACAGT 955
QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGluLeu 319
DB 956 AGCCGCTTGCAATCCAGGACAGTGGCCAGATCCACCTTGGCAGATTCCTCCGAGCTC 1015
QY 320 LeuSerAspSerSerAsnSerAsnCysIleThrTyrGlnGlyThrAsnGlyLysIlePheLys 339
DB 1016 CTGTGCGACAGCTCCAACTCCAGCTGATACCTGGGAAGGCAACAGGGGAGTTCAAG 1075
QY 340 MetThrAspProAspGluValAlaArgArgTyrPglValArgLysSerLysProAsnMet 359
DB 1076 ATACAGGATCCGACAGCGGTGGCCCGCGCTGGGAGAGCGGAAGCAACCAACATG 1135
QY 360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
DB 1136 AACTTACGATTAAGTCCAGCCGCGCTCCGTTACTATATGACAAAGACATCAAGACCAAG 1195
QY 380 ValHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399

```

```

DB 1196 GTC----- 1198
QY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
DB 1199 ---CATGGGAAGCGCTACGGCTCAAGTTCGACTTCCACGGAGATCGCCAGCGCTCAG 1255
QY 420 Pro----- 420
DB 1256 CCCCACCCCGGAGTCACTCTGTACAGTACCCCTCAGACCTCCGTAATGAGGCTCC 1315
QY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439
DB 1316 TATCAGCCGCCACCCACAGAAATGATGACTTGTGGCGCCCAACCTCCAGCTCCCGTG 1375
QY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyIle 459
DB 1376 ACATCTTCCAGTTTGTGTGCTGCCCAACCCCTACTGGAATTCACCAACTGGGGGTATA 1435
QY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
DB 1436 TACCCCAACACTAGGCTCCACACGCAATATGCTTCTCATCTGGGCACTACTAC 1492

RESULT 7
US-09-726-811-4920
; Sequence 4920, Application US/09726811
; GENERAL INFORMATION:
; APPLICANT: Gutierrez-Ramos, Jose-Carlos
; APPLICANT: Welch, Nadine S.
; APPLICANT: Wen, Danyl
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2027-001
; CURRENT APPLICATION NUMBER: US/09/726, 811
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/168, 136
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 5515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4920
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-811-4920

Alignment Scores:
Pred. No.: 4,4e-159 Length: 3309
Score: 2195.50 Matches: 416
Percent Similarity: 84.97% Conservative: 8
Best Local Similarity: 83.37% Mismatches: 10
Query Match: 84.83% Indels: 65
DB: 29 Gaps: 4

US-09-902-772-4 (1-478) x US-09-726-811-4920 (1-3309)
QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerLysLysGlnSerLeuPheGlu 20
DB 128 ATGGCCAGCAGCTATTAAGGAAGCTTATCGATTGTGATGAGAGACAGTGTGTTGAG 187
QY 21 CysAlaTyrGlySerProHISLeuAlaLysThrGluMetThrAlaSerSerSerglu 40
DB 188 TGTGCTACGGAAGCCACACCTGGCTTAACAGAGATGACCCGCTCTCTCCAGCGAC 247
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerglnPro 60
DB 248 TATGAGACAGACTTCCAAAGATGAGAGCCAGCGCTCCCTCAGCAGATGGCTGTCTCAACC 307
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 308 CCAAGCCAGGGTCCACCAAAATGGAATGTACCTTACCCAGCGTGAATGGCTCAAGGAC 367
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100

```

```

Db      368  TCCTCGATGATGACAGTGTGGCCAAAGCGGGAGATGGTGGCAGACCCAGACACCGTT  427
Qy      101  GYMetasntyrGlySerTyrMetGluGluHisIleProProProAsnMetThrThr  120
Db      428  GGGATGATACAGCGCAGCTACATGAGAGAGAGACAGTGCACCCCAACATATACACG  487
Qy      121  AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg  140
Db      488  AACGAGCGCAGAGTTATCGTGCAGCAGATCTCTACGCTATGAGATACAGACATGTGCG  547
Qy      141  GluTrpLeuGluTrpAlaValAlaValGluTyrGlyLeuProAspValAspIleLeuPhe  160
Db      548  CAGTGGCTGGAGTGGGGGTGAAGAATATGGCTTCACACGCAACATCTGTATTC  607
Qy      161  GluAsnIleAspGlyGlyGluLeuGlyMetThrTyrAspAspPheGluArgLeuThr  180
Db      608  CAGACATCGATGGGAAGAGACGTGTGCAATGACCAAGAGACGCTTCAGAGGCTCAC  667
Qy      181  ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluThrProLeu  200
Db      668  CCCAGCTACACGCGCAGCTCTCTCTCAGATCTCCACTACAGAGAGACTCTCT  727
Qy      201  ProHisLeuThrSerAspAspValAspValAlaLeuGluAsnSerProAlaGlyLeuHis  220
Db      728  CCACATTTGACTTCAGATGATGTTGATTAAGCCTTACAAAACCTCCACGCTTATGAT  787
Qy      221  AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla  240
Db      788  GCTAAGAAACACA-----  799
Qy      241  ThrGluArgIleThrThrArgProAspLeuProTyrGluGluAlaArgArgSerAlaTyr  260
Db      800  -----GATTTACCATATAGAGCCCGCCAGAGATACAGCTCG  835
Qy      261  ThrSerHisSerHisProThr---GlnSerTyrAlaThrGlnProSerSerSerThrVal  279
Db      836  ACCGGTACGCGCCACCCACGCGCCCACTGCAAAAGCTCTCAACATCTCTCCACAGTG  895
Qy      280  ProLysThrGluAspGluArgProGluGlnLeuAspProTyrGlnIleLeuGlyProThrSer  299
Db      896  CCCAAACTGTAAGACGCGCTCTCATGTTAGATCCTTATCAGATTTCTGGACCAACAGT  955
Qy      300  SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeu  319
Db      956  AGCGGCTTGCAATCCAGGACAGTGGCAGATCCAGGTTTGCGAGTCTCTCGAGCTC  1015
Qy      320  LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyGluPheLys  339
Db      1016  CTCTCGACAGCTCCAACTCAGCTGATCAGCTGGGAAGGCAACAGGGGAGTTCAAG  1075
Qy      340  MetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet  359
Db      1076  ATGACGAGATCCGACAGAGTGGCCCGGCTGGGAGAGGGAAGCAACCAACCAATG  1135
Qy      360  AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys  379
Db      1136  AACTAGATTAAGTCCAGCGCGCTCCGTTACTACTATGACAAGAACATCATGACCAAG  1195
Qy      380  ValHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer  399
Db      1196  GTC-----  1198
Qy      400  TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln  419
Db      1199  ---CATGGAAAGCGCTACAGCTCAAGTTGACTCCACGAGGATCGCCGAGCGCTTCAG  1255
Qy      420  Pro-----  420
Db      1256  CCCCACCGCCGAGTCACTCTGTACAAAGTACCCCTCAGACCTCCCGTACATGGGCTCC  1315
Qy      421  ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal  439
Db      1316  TATCAGCGCCACCAAGAAAGATGAACTTTGTGGCGCCCAACCTCAGCGCTCCCGG  1375

```

```

Qy      440  ThrSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIle  459
Db      1376  ACATCTCCAGTTTGTGGTCCCAACCAACCAATGAGATTCACCACTGGGGGTATA  1435
Qy      460  TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr  478
Db      1436  TACCCCAACACTAGGCTCCCGCAGCAGCATATGCTTCTCATCTGGGCGCTACTAC  1492

RESULT 8
US-09-948-941-233
; Sequence 233, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1000788
; CURRENT APPLICATION NUMBER: US/09/948, 941
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; NUMBER OF SEQ. ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-233

Alignment Scores:
Pred. No.: 2,22e-158 Length: 1389
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: Gaps: 4

US-09-902-772-4 (1-478) x US-09-948-941-233 (1-1389)
Qy      2  AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys  21
Db      25  GCAGCTATATCAAGAGAGCGCTTATCATGTTGATGAGAGACCAAGTGTGTTGACTGT  84
Qy      22  AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr  41
Db      85  GCCTACGGAAGCGCACACCTGGCTTAAGACAGATGACCCGCTCTCTCCAGCACTAT  144
Qy      42  GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro  61
Db      145  GGACAGACTTCCAAAGATGAGCCCGGCTCCACAGGATGGCTGCTCAACCCCA  204
Qy      62  AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer  81
Db      205  GCCAGGCTCACATCAAAATGAATGATTAACCTTACCCAGGTGAATGGCTCAAGAACTCT  264
Qy      82  ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGly  101
Db      265  CTTGATGAATAGCAGTGTGGCCAAAGCGGGAAGATGGTGGCGCCAGACACCGTTGGG  324
Qy      102  MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrTrpAsn  121
Db      325  ATGAACCTACGCGCAGCTPACATGAGAGAGCAATCCACCCCAACATGACACCAAG  384
Qy      122  GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln  141
Db      385  GAGCGCAGATTTCTGCTGCACAGCAGATCCAGCTATGAGTACAGACATGTGGCGCAG  444
Qy      142  TrpLeuGluTrpAlaValLysGlyTyrGlyLeuProAspValAspIleLeuLeuPheGln  161
Db      445  TGGCTGAGAGTGGCGGTGAAGAATATGGCTTCCAGAGCTCAACATCTTGTATTCAG  504
Qy      162  AsnIleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheIleArgLeuThrPro  181

```

Db 505 AACATCATGGGAAGAACTGTGCAAGATGACCAAGACGACTTCACAGAGCTCACCCCC 564  
 Qy 182 SerTyraSnAlaaspIleLeuSerHisIleuHisTyrlleuArgIuThProLeuPro 201  
 Db 565 AGCTACACAGCCGACATCTCTCTCATCTCCATCTCCATCCACAGAGAGACTCTCTTCCA 624  
 Qy 202 HisLeuThrSerAspAspValAspLysAlaIleuGlnInsSerProAlaGluMetHisAla 221  
 Db 625 CATTTGACTTCAGATGATGATGATTAAGCCCTTACAAAACCTCCACGGTTAATCATAGCT 684  
 Qy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241  
 Db 685 AGAAACACACA----- 693  
 Qy 242 GlnArgIleThrThrArgProAspLeuProGlyGlnGlnAlaArgSerAlaTyrThr 261  
 Db 694 -----GATTTACCATATGAGCCCCCAAGAGATGAGCTTGAGCC 732  
 Qy 262 SerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThrValPro 280  
 Db 733 GGTACAGGCCACCCACGCCCCACATCGAAGACCTGCTACACATCTCTCTCCACAGTGCCTC 792  
 Qy 281 LysThrGluAspGlnArgProGlnLeuAspProGlyGlnIleuGlyProThrSerSer 300  
 Db 793 AAAACTGAGACACAGCGCTCCATGATCTTATCATGATTTCTTGAGACCAACAAGTAGC 852  
 Qy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheIleuLeuGlnLeu 320  
 Db 853 CGCCTTGCAATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 912  
 Qy 321 SerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLysPheLysMet 340  
 Db 913 TCGGACACGCTCCACATCCAGTGCATCACCTGGGAGACCAAGGAGGAGGAGGAGGAGGAGG 972  
 Qy 341 ThrAspProAspGluValAlaIleArgTrpGlyLysLysSerLysProAsnMetAsn 360  
 Db 973 ACGGATCCCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1032  
 Qy 361 TyrAspLysLeuSerArgAlaLeuArgTrpTyrTyrAspLysAsnIleMetThrLysVal 380  
 Db 1033 TACATTAAGCTACACCGGCGCTCCGTTACTACTATGACAAAGAACATCATGACCAAGGTC 1092  
 Qy 381 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400  
 Db 1092 ----- 1092  
 Qy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420  
 Db 1093 CATGGGAAGCGCTACGCTCAAGTTCGACTTCCACGAGGAGTCCGCCAGGCCCTCCAGGCC 1152  
 Qy 420 ----- 420  
 Db 1153 CACCCCGGAGATCTGTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCTAT 1212  
 Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440  
 Db 1213 CAGCGCCACCCACGAGAGATGAACTTGTGGCGCCACACCTCCAGGCCCTCCCGGTGACA 1272  
 Qy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 460  
 Db 1273 TCTTCCAGTTTGTGTGGCCCAACCCATCTGGAATTCACCACTGGGGGTATTTATC 1332  
 Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478  
 Db 1333 CCCAACACTAGGCTCCCAACGACCATATGCTTCTCATCTGGGCGACTTACTAC 1386  
 RESULT 9  
 PCT-US02-23913-98  
 : Sequence 98, Application PC/TUS0223913  
 : GENERAL INFORMATION:  
 : APPLICANT: Millennium Pharmaceuticals, Inc. et al.  
 : TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
 : TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

: TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
 : FILE REFERENCE: MRI-044PC  
 : CURRENT APPLICATION NUMBER: PCT/US02/23913  
 : CURRENT FILING DATE: 2002-07-25  
 : PRIOR APPLICATION NUMBER: 60/307,982  
 : PRIOR FILING DATE: 2001-07-25  
 : PRIOR APPLICATION NUMBER: 60/314,356  
 : PRIOR FILING DATE: 2001-08-22  
 : PRIOR APPLICATION NUMBER: 60/325,020  
 : PRIOR FILING DATE: 2001-09-25  
 : PRIOR APPLICATION NUMBER: 60/341,746  
 : PRIOR FILING DATE: 2001-12-12  
 : PRIOR APPLICATION NUMBER: 60/362,158  
 : PRIOR FILING DATE: 2002-03-05  
 : NUMBER OF SEQ ID NOS: 455  
 : SOFTWARE: FASTSEQ for Windows Version 4.0  
 : SEQ ID NO 98  
 : LENGTH: 3166  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 PCT-US02-23913-98  
 Alignment Scores:  
 Pred. No.: 6,02e-158 Length: 3166  
 Score: 2180.50 Matches: 413  
 Percent Similarity: 84.74% Conservative: 9  
 Best Local Similarity: 82.93% Mismatches: 11  
 Query Match: 84.25% Indels: 65  
 DB: 1 Gaps: 4  
 US-09-902-772-4 (1-478) x PCT-US02-23913-98 (1-3166)  
 Qy 2 AlaSerThrIleLysGluAlaLeuSerValSerGluAspGlnSerLeuPheGluLys 21  
 Db 281 GCACCTATATCAAGAGAAAGCTTATCAGTTGTGAGTAGAGACCAAGTGTGTGAGTGT 340  
 Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlyTyr 41  
 Db 341 GCCTACGGAAGCCACACACTGGCTTAAGACAGATGACAGCGGCTCTCTCCAGGACTAT 400  
 Qy 42 GlyIleThrSerLysMetSerProArgValProGlnGlnAspTyrPheLysGlnProPro 61  
 Db 401 GGACAGACTTCCAGATGAGAGCCACGCGTCCCTCAGCAGATGAGTGTGCTCAACCCCA 460  
 Qy 62 AlaArgValThrIleLysMetGluLysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
 Db 461 GCCAGGGTCCACATCAAAATGGAATGTAACCTTACGCCAGGTAATGCTCAAGAACTCT 520  
 Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnValGly 101  
 Db 521 CTTATGATATGCACTGTGGCCCAAGGCGGAGATGGTGGCAGCCACCCAGACCCGTTGGG 580  
 Qy 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121  
 Db 581 ATGAATACAGGACACTCATGAGAGAGAGACATGACACCCCAAAACATGACCAACAGMAC 640  
 Qy 122 GlnArgArgValIleValProAlaAspProThrIleuThrPheSerThrAspHisValArgGln 141  
 Db 641 GAGGCGAGAGTTATCGGCCAGCAGATCTCAAGCTATGAGATGACAGCAATGTCGGCAG 700  
 Qy 142 TrpLeuAluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161  
 Db 701 TGGCTGAGTGGGCGGAGGAGAAATATAGGCTTCCAGACGTCACATCTGTATTTCAG 760  
 Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181  
 Db 761 AACATCGATGGAGGAAGAACTGTCAAGATGACCAAGAGACGACTCCAGAGGCTCACCC 820  
 Qy 182 SerTyraSnAlaaspIleLeuSerHisIleuHisTyrlleuArgIuThProLeuPro 201  
 Db 821 AGCTACACAGCGCGACATCTTCTCTCATCTCCATCTCCACAGAGAGACTCTCTTCCA 880  
 Qy 202 HisLeuThrSerAspAspValAspLysAlaIleuGlnInsSerProAlaGluMetHisAla 221



```

|||||
Db 881 CATTGACTCAGATGATGTGATAAAGCCTTACAAAACCTCCACGGTAAATGACATGCT 940
Qy 222 ArgSnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThr 241
Db 941 AGAACAACA----- 949
Qy 242 GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTyrThr 261
Db 950 -----GATTACATATAGAGCCGCCAGAGATGACCTGGAGCC 988
Qy 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 989 GGTACAGCGCCACCCACGCGCCACGAAAGCTCCTCAACCATCTCCTTCCACAGATGCC 1048
Qy 281 LysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db 1049 AAAACTGAAGACGAGCGCTCCATGATCCATTATGATGATTTGGACCAACAAGTAGC 1108
Qy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeu 320
Db 1109 CGCTTGCAATCCAGGCAAGTGGCCAGATCCAGCTTGGCAGTTCCTCTGAGACTCTCG 1168
Qy 321 SerAspSerSerAsnSerAsnCysIleThrTyrGluGlyThrAsnGlyLuphLysMet 340
Db 1169 TCGGACAGCTCCACATCCAGCTCAGTCATCAGTGGGAAGGACCAACGGGAGTTCAGATG 1228
Qy 341 ThrAspProAspGluValAlaArgArgTyrGlyLuphLysSerLysProAsnMetAsn 360
Db 1229 AGGATCCCGACGAGGAGGCGCGGCGGAGAGGAGGAGCAACCAACATGATGAC 1288
Qy 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db 1289 TAGCATAGCTACACCGCGGCGCTTACTACTATAGCAAGAACATCATACCAAGTGC 1348
Qy 381 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1348 ----- 1348
Qy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1349 CATGGAGAGCGCTACGCTTACAGTTCACACGCGGATCGCCAGGCCCTCCACGCC 1408
Qy 420 ----- 420
Db 1409 CACCCCGGAGTATCTGTACAAAGTACCCCTCAGACCTCCGTCATATGCGCTCTAT 1468
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1469 CAGCCCGACCCACGAGAGATGAATTTGTGGCGCCACCCCTCCAGGCCCTCCCGTGACA 1528
Qy 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyGlyIleTyr 460
Db 1529 TCTTCCAGTTTTTTGTGTCGCCCAACCCATACGGAATTCACCAACTGGGGGATATATAC 1588
Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1589 CCCAACACTAGGCTCCCGACAGCATATGCTTCATCTGGGCACTTACTAC 1642

```

```

; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-784-356-26

Alignment Scores:
Pred. No.: 6,02e-158 Length: 3166
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 30 Gaps: 4

US-09-902-772-4 (1-478) x US-09-784-356-26 (1-3166)
Qy 2 AlAserThrIleLysGlnAlaLeuSerValValSerLupAspGlnSerLeuPheGluCys 21
Db 281 GCACCTCATATCAAGGAAGCCTTATCAGTTGTGATGAGACCCAGTGTGTTGAGTGT 340
Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41
Db 341 GCTTACGGAAGCGCCACCTGCTGCTAAGACAGATGACCGGTCTCTCCAGGACTAT 400
Qy 42 GlynThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnProPro 61
Db 401 GGACAGACTTCCAAAGATGAGAGCCAGCGTCCCTCAGCAGATGGTGGCTCAACCCCA 460
Qy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerThrArgSer 81
Db 461 GCCAGGCTCACCAACAAATGGAATGTAAACCTTACCCAGGTGAATGGCTCAAGCACTCT 520
Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db 521 CCGTATGAATGCACTGTGGCCAAAGCGGGAAGATGGTGGACACCCAGACCGGTGGG 580
Qy 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121
Db 581 ATGAATCAACGGCAGCTCATGTGAGAGAACACATGCCCCCAACATGACACGAC 640
Qy 122 GlnArgArgValIleValProAlaAspProThrLeuThrPheThrAspHisValArgGln 141
Db 641 GAGCGAGAGTATATCGCCAGCAGATCTTACGCTATGAGATACAGACCATGTGCGGCGAG 700
Qy 142 TyrLeuGlnTyrPalaValLysGluTyrGlyLeuPProAspValAspIleLeuPheGln 161
Db 701 TGGCTGAGTGGGGGGGGAAGATATATGGCTTCCAGACGTCAACATCTTGTATTCCAG 760
Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 761 AACATCGATGGGAAGAACTGTGCAGATGACCAAGAGACGACTCCAGAGGCTCACCCCC 820
Qy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro 201
Db 821 AGCTACAAACCCGACATCTTCTCCTCACATCTCCACATCTCCAGAGAGATCTCTTCCA 880
Qy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db 881 CATTGACTTCAGATGATGTGATTAAGCTTACAAAACCTCCACAGGTTAAAGCATGCT 940
Qy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThr 241
Db 941 AGAACAACA----- 949
Qy 242 GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTyrThr 261
Db 950 -----GATTACATATAGAGCCGCCAGAGATGACCTGGAGCC 988

```





```
|||||
Db 1109 CGCCTTGCATTCAGGAGGAGTGGCCAGATCCAGCTTGGCGAGTTCCTCGAGCTCTG 1168
Qy 321 SerSpSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 340
Db 1169 TCGGACAGCTCCAACTCCAGTGCATCAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1228
Qy 341 ThrSpProAspGluValAlaArgArgTrpGlyLysSerSerSerSerSerSerSerSerSer 360
Db 1229 ACGGATCCCGACGAGGAGTGGCGCGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1288
Qy 361 TyrAspLysLeuSerArgAlaLeuArgTrpTrpTrpTrpTrpTrpTrpTrpTrpTrpTrp 380
Db 1289 TACGATAGCTCAGCCCGCGCTCCGTTACTATATAGACAGACATCATGACCAAGGTC 1348
Qy 381 HisProProGluSerSerSerMetTyrLysTyrProSerSerSerProTyrMetSerSerTyr 400
Db 1348 ----- 1348
Qy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyLeuAlaGlnAlaLeuGlnPro 420
Db 1349 CATGGAGAGCGCTACGCTACAAAGTTGAGCTTCACAGGAGATCGGCCAGGCCCTTCAGGCC 1408
Qy 420 ----- 420
Db 1409 CACCCCGGAGTATCTGTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCTAT 1468
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1469 CAGCCCGACCCACGAGAGATGACTTGTGGCGCCCAACCTCCAGCCCTCCCGCTGACA 1528
Qy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysTyr 460
Db 1529 TCTTCAGTGTGTTTGTGTCGCCCAACCCATATGGAATTCCACAGCTGGGGGTATATAC 1588
Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyTrpTyrTyr 478
Db 1589 CCCAACACTAGCTCCGCCACGACCATATGCTTCTCATCTGAGGCGACTTACTAC 1642

RESULT 12
US-10-205-823-98
: Sequence 98, Application US/10205823
: GENERAL INFORMATION:
: APPLICANT: Schlegel, Robert
: APPLICANT: Monahan, John E.
: APPLICANT: Endege, Wilson O.
: APPLICANT: Gannavarapu, Manjula
: APPLICANT: Gorbacheva, Bella
: APPLICANT: Hoersch, Sebastian
: APPLICANT: Kamatkar, Shubhangl
: APPLICANT: Monsey, Angela M.
: APPLICANT: Gialt, Karen
: APPLICANT: Zhao, Xumei
: APPLICANT: Anderson, Dustin
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
: TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: FILE REFERENCE: MRI-044
: CURRENT APPLICATION NUMBER: US/10/205, 823
: PRIOR APPLICATION NUMBER: 2002-07-25
: PRIOR FILING DATE: 2001-07-25
: PRIOR APPLICATION NUMBER: 60/314,356
: PRIOR FILING DATE: 2001-08-22
: PRIOR APPLICATION NUMBER: 60/325,020
: PRIOR FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: 60/341,746
: PRIOR FILING DATE: 2001-12-12
: PRIOR APPLICATION NUMBER: 60/362,158
: PRIOR FILING DATE: 2002-03-05
: NUMBER OF SEQ ID NOS: 455
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 98
```

```

: LENGTH: 3166
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-205-823-98

Alignment Scores:
Pred. No.: 6,02e-158 Length: 3166
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 42 Gaps: 4

US-09-902-772-4 (1-478) x US-10-205-823-98 (1-3166)

Qy 2 AlasSerThrIleLysGluAlaLeuSerValValSerLysLysAspGlnSerLeuPheGluCys 21
Db 281 GCAGCTCATATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 340
Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerSerGluTyr 41
Db 341 GCTTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
Qy 42 GlyLysThrSerLysMetSerProArgValProGlnLysAspThrLeuSerGlnProPro 61
Db 401 GGACAGAGCTTCCAGAGATGAGCCCGCGCTCCAGAGAGATGGCTGTCAACCCCA 460
Qy 62 AlaArgValThrIleLysMetGluCysAsnProGlnGlnAlaSerSerSerSerSerArgAsnSer 81
Db 461 GCCAGGCTCACCATCAAAATGAAATGATGTAACCTTACCCAGGTGAATGGCTCAAGGAGACTCT 520
Qy 82 ProAspCysSerValAlaLysGlyLysMetValSerSerSerSerSerSerSerValGly 101
Db 521 CCTGATGAATGACAGTGTGGCCAAAGCGGGAGAGATGGTGGGAGCCAGACCGCTTGGG 580
Qy 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121
Db 581 ATGAACATACGCGACGCTCATGAGAGAGACACATGCGACCCCAAAACATGAGACCCAGC 640
Qy 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db 641 GAGCGCAGAGTATTCGTCGCCAGCAGATCTTACGAGTACAGTACAGGAGGAGGAGGAGGAGGAG 700
Qy 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db 701 TGGCTGAGTGGGGGCGTGAAGAATATGAGCTTCCAGACGTCACATCTGTTATTCAG 760
Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 761 AACATCGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 820
Qy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluTrpProLeuPro 201
Db 821 AGCTACAGCCCGACATCTTCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 880
Qy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db 881 CATTTGACTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 940
Qy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db 941 AGAAACACA----- 949
Qy 242 GluArgLeuThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThr 261
Db 950 -----GATTTACCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 988
Qy 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerSerThrValPro 280
Db 989 GGTACAGGCGACCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1048
Qy 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db -----
```

```

Db      1049 AAAATGAAACGACGCTCTCAGTAGATCCCTTATCATGATCTTGGACCAACAGTAGC 1108
Qy      301  ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPglInPheLeuGlnLeu 320
Db      1109 CGCCTTGCAAAATCCAGGAGTGCGCAGATCCAGCTTGGCAGTTCTCCCTGGACCTCTG 1168
Qy      321  SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMet 340
Db      1169 TCGGACAGCTCCAACTCCAGTCAGTCATCACCGGGGAAGGACCAACGGGGAGTTCAAGATG 1228
Qy      341  ThrAspProAspGluValAlaIaArgTrpGlyGluArgLysSerLysProAsnMetAsn 360
Db      1229 ACGGATCCCGACGAGGAGTGCGCCGCGCTGGGAGAGCGGAAGCAACCAACCAATGTAAC 1288
Qy      361  TyrAspLysLeuSerArgAlaLeuArgTrpTyrTrpLysPheLysAsnIleMetThrLysVal 380
Db      1289 TACATATAGCTCAGACCGCGCCCTCGTACTACTATGACAAAGAACATCATATCACAAGGTC 1348
Qy      381  HisProProGluSerSerMetLysTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db      1348 ----- 1348
Qy      401  HisGlyLysArgTrpAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1349 CATGGAGAGCGCTACGGCTTACAAAGTGCAGCTCCACGGGAGATCCGCCAGGCCCTCCAGCCC 1408
Qy      420 ----- 420
Db      1409 CACCCCGGAGATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGGCTCTAT 1468
Qy      421  HisAlaHisProGlnLysMetAsnPheValAlaIaProHisProProAlaLeuProValThr 440
Db      1469 CAGCCACACCCAGAGATGAATCTTGTGGCGCCACCCCTCCAGGCCCTCCCGGTACGA 1528
Qy      441  SerSerSerPhePheAlaIaIaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 460
Db      1529 TCTTCCAGTTTCTTGTGTCGCCCAACCCATCTGGAATTCACCACTGGGGGTATATATAC 1588
Qy      461  ProAsnThrArgLeuProAlaIaHisMetProSerHisLeuGlyThrTyr 478
Db      1589 CCCAACACTAGGCTCCCCACACGACCATATGCTTCTCATCTGGGCACTTACTAC 1642

RESULT 13
US-60-207-355
; Sequence 355, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: C001321
; CURRENT APPLICATION NUMBER: US/60/360,207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 355
; LENGTH: 2172
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-355

Alignment Scores:
Pred. No.: 1,52e-145 Length: 2172
Score: 2017.50 Matches: 384
Percent Similarity: 79.12% Conservative: 10
Best Local Similarity: 77.11% Mismatches: 7
Query Match: 77.96% Indels: 97
DB: 80 Gaps: 3

US-09-902-772-4 (1-478) x US-60-360-207-355 (1-2172)
Qy      1  MetaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db      107  ATGGCCACGACATAATTAAAGAGGCTTGTCACTTGTGACGAGGACCAAGTACATAATTGAG 166
Qy      21  CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40

```

```

Db      167  TGTGCTACGGAAGCCACACCTGGCTAAAGACAGATGACCCGATCCTTCCAGTGAC 226
Qy      41  TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
Db      227  TATGGCCAGACATCAAGATGAGTCCAGAGATCTCCCTCAGCGAGCTGGCTTCTTCAAGCC 286
Qy      61  ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db      287  CCAGCCAGGCTCACCATCAATGATGAGTGCACACCTTGTACGTTGATGTTCCACAGAAC 346
Qy      81  SerProAspPheCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db      347  TCACCTGATGATGCAGTGAACAAAGGTGGAAAGATGTTGGGACAGCCGAGTACAGTGTG 406
Qy      101  GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
Db      407  GGGATGAGCTACGGCAGCTACATGAGGAGAGCATGTGCCCTCCCAATATACACACA 466
Qy      121  AsnGluArgArgValIleValProAlaAspProThrLeuThrPheSerThrAspHisValArg 140
Db      467  AATAGCGCAGAGATGATCTCTCCAGATCTTCTGTGGACACACAGACCATATGTCGA 526
Qy      141  GlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
Db      527  CAGTGGCTGGAGTGGGGGTGAAGAATATGCGCTCCGATGTGAGAGCTTACTATTT 586
Qy      161  GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db      587  CAGAAATATCGATGGAGGAGGAGCTGTGCAGATGTCAAAAGATGACTTCCAGCGGCTCAGC 646
Qy      181  ProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
Db      647  CCGAGCTACAAATGCCGACATTTCTTCTACATCTCCACATCTCCAGAGACATCCCTT 706
Qy      201  ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
Db      707  CCACATCTGACTTCCGATGAGCTGTGATTAAGCTTTTACAAAACCTCCACAGGTTAATGAT 766
Qy      221  AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
Db      767  GCCAGAAACACAGGGGGGTGCGAGCTTTATTTTCCAAATATTGATGATATATCCCAAGCT 826
Qy      241  ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 260
Db      827  ACGCAAAGAATTACAACTAGG----- 847
Qy      261  ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db      847 ----- 847
Qy      281  LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db      847 ----- 847
Qy      301  ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPglInPheLeuGlnLeu 320
Db      848 ----- 895
Qy      321  SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMet 340
Db      896  TCGAGACAGCTCCAACTCCAGTCAGTCACCGGGGAAGGACCAACGGGGAGTTCAAGATG 955
Qy      341  ThrAspProAspGluValAlaIaArgTrpGlyGluArgLysSerLysProAsnMetAsn 360
Db      956  ACAAGCCCGACGAGGAGTGTGCGCGCTGGGGGAGAGGAAGACCAACCAATCATTAAC 1015
Qy      361  TyrAspLysLeuSerArgAlaLeuArgTrpTyrTrpLysPheLysAsnIleMetThrLysVal 380
Db      1016  TATGACAGAGCTCAGCCCGCCCTCGCTACTACTACACAAACAAATCATATGACCAAGGTG 1075
Qy      381  HisProProGluSerSerMetLysTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400

```

```

Db 1075 ----- 1075
Qy 401 HisGlyLysArgTyrAlaTyrLysPheaspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1076 CACGGGAAGCGCTACGCTACAGTTGACTTCCACGGGATGGCCAGGCCCTGACGCC 1135
Qy 420 ----- 420
Db 1136 CACCTCTGAGTGCCTTACAGTACCCTCCGACCTGCATACATGAGGCTCTAT 1195
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1196 CACGGCCACCCCGAGAGATGACTTGTGTCTCCCGACCTCCCGCTCCACGTACA 1255
Qy 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGlyIleTyr 460
Db 1256 TCTTCACAGTTCTTGTGCTTCCCGAACCCATACGTGATTCACCGACGAGGGGATCTAC 1315
Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
Db 1316 CCGAACACTAGGCTCCCGACGACCATATGCTCTCCTCAGTGGCAGCTACTAC 1369

```

## RESULT 14

```

US-09-442-384A-475
; Sequence 475, Application US/09442384A
; GENERAL INFORMATION:
; APPLICANT: Lukashiev, Matvey
; TITLE OF INVENTION: Hematology/Immunology Array
; FILE REFERENCE: CLON-006CIP15
; CURRENT APPLICATION NUMBER: US/09/442, 384A
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 09/053, 375
; NUMBER OF SEQ ID NOS: 830
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 3126
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-442-384A-475

```

```

Alignment Scores:
Pred. No.: 3.4e-129 Length: 3126
Score: 1808.50 Matches: 342
Percent Similarity: 82.23% Conservative: 5
Best Local Similarity: 81.04% Mismatches: 10
Query Match: 69.88% Indels: 65
Gaps: 4

```

US-09-902-772-4 (1-478) x US-09-442-384A-475 (1-3126)

```

Qy 78 SerArgAsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSer 97
Db 150 TCAAGGAACCTCTCTGATGATGCAATGTCGCCAAGCGGGAAGATGTCGACGCCA 209
Qy 98 AspAsnValGlyMetAsnTyrGlySerTyrMetGlnLysHisIleProProAsn 117
Db 210 GACACCGCTGGATGACACTACGACGACATGAGAGAGAACACATGCCACCCCAAC 269
Qy 118 MetThrThrAsnGlnArgValIleValProAlaAspProThrLeuTyrPasnSerThr 137
Db 270 ATGACCCAGAACGCGCCAGATGATTCGTGCCACCAATCTTACGCTATGAGTACAC 329
Qy 138 HisValArgGlnTyrPasnGlnTyrPalaValLysGlnTyrGlyLeuProAspValAsp 157
Db 330 CATGTGGCGCAGTGTGTCGATGCGCGTGAAGAAATATGCGCTCCAGACGTAAACATC 389
Qy 158 LeuLeuPheGlnAsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGln 177
Db 390 TTGTTATTCAGAACATCATGATGGAGAACACTGTCAAGATGACCAAGACGACTTCAG 449
Qy 178 ArgLeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGln 197

```

```

Db 450 AGGCTACACCCCGACCTACACGCGACATCCTTCTCACATCTCCACTACCTACGAGAG 509
Qy 198 ThrProLeuProHisLeuThrSerAspAspValAspLysAlaGlnAsnSerProArg 217
Db 510 ACTCTCTCCACATTTGACTTCAGATGATGTATTAAGCTTACAAACTCTCCACGG 569
Qy 218 LeuMetHisAlaArgAsnThrGlyIleAlaThrPheIlePheProAsnThrSerValTyr 237
Db 570 TTAATGATGCTATAAACACA ----- 590
Qy 238 ProGlnAlaThrGlnAlaGlyIleThrThrArgProAspLeuProTyrGlnAlaArg 257
Db 591 ----- GATTACCATATGAGCCCGCCAGAGA 617
Qy 258 SerAlaThrThrSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSer 276
Db 618 TCAGCTGGACCGGTCTACGCGCCACCGCCAGTCGAAAGCTGTCAACTCTCTCT 677
Qy 277 SerThrValProLysThrGlnArgProGlnLeuAspProTyrGlnIleLeuGly 296
Db 678 TCCACAGTGGCCAAACTGAAGACACGCTCTCAGTTAGATCCTTATCAGATTCTTGA 737
Qy 297 ProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeu 316
Db 738 CCAACAAATAGCCCGCTTGCAGAAATCCAGCGAGTGCACAGATCCAGCTTGTGCACTTCTCTC 797
Qy 317 LeuGlnLeuLeuSerAspSerSerAsnSerAsnLysIleThrThrProGlnGlyThrAsnGly 336
Db 798 CTGGAGCTCTGTGGACGCTCCAGCTCCAGCTGATCAGCTGGGAAGGACCCAAAGGG 857
Qy 337 GluPheLysMetThrAspProAspGlnValAlaArgArgTyrGlyGlnArgLysSerLys 356
Db 858 GAGTTCAAGATGACGAGATCCCGAGAGAGTGGCCCGGCGGAGAGCGGAAGAGCAA 917
Qy 357 ProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrAspLysAsnIle 376
Db 918 CCCAACATGAACTGATAGCTCAGCGCGCCCTCGTTCATCACTATGACAAAGAACATC 977
Qy 377 MetThrLysValHisProGlnLysSerMetLysTyrProSerAspLeuProTyr 396
Db 978 ATGACCAAGGTC ----- 989
Qy 397 MetSerSerTyrHisGlyLysArgTyrAlaTyrLysPheaspPheHisGlyIleAlaGln 416
Db 990 -----CATGGGAAGCGCTACGCTACAAAGTTGACACTTCCACGGGATCGCCAG 1037
Qy 417 AlaLeuGlnPro ----- 420
Db 1038 GCCCTCCAGCCCGACCCCGGAGTCAATCTGTACAGTACCCTTCAGACTCCCGTAC 1097
Qy 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAla 436
Db 1098 ATGGGCTCTATACAGCGCCACCCAGACAAATGATGACTTGTGGCGCCCGCCCTCCAGCC 1157
Qy 437 LeuProValThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThr 456
Db 1158 CTCGCCGTGACATCTTCCAGTTTGTGTCGCCCAAAACCATATGATTCAGTTCACCAACT 1217
Qy 457 GlyGlyIleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThr 476
Db 1218 GGGGGTATATACCCCAACTAGGCTCCACACGACCATATGCTCTCAATGCGGCACT 1277
Qy 477 TyrTyr 478
Db 1278 TACTAC 1283

```

## RESULT 15

```

US-09-442-384B-475
; Sequence 475, Application US/09442384B
; GENERAL INFORMATION:
; APPLICANT: Lukashiev, Matvey

```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:30:39 ; Search time 1092.35 Seconds

(Without alignments)  
3233.216 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKFALSVEQSLFE.....IYPNTRLPAAHMSHLGTYT 478

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Ygapext 0.5	
Xgapop 6.0 , Ygapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 8407786 segs, 3694357880 residues

Total number of hits satisfying chosen parameters: 16815572

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+2pn.model -DEV=xlh  
-Q=/cgn2\_1/USPto.spool/US0902772/rnat.23072003\_093658\_14943/app\_query.fasta.1.1294  
-DB=Pending\_Patents\_NA\_New -QFMT=fastap -SUFFIX=trnp -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -FRANS=human40.cdl  
-LIST=45 -DOCALLIG=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US0902772\_6C8N\_1.1.325\_6runat.23072003\_093658\_14943 -NCPu=6 -ICPU=3  
-NOMMR -LARGEOUTRT -NEG.SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending\_Patents\_NA\_New:\*  
2: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:\*  
8: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:\*  
9: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:\*  
10: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:\*  
11: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:\*  
12: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:\*  
13: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*  
14: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*  
15: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*  
16: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*  
17: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*  
18: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*  
19: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*  
20: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

# SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	2324.5	89.8	1930	15	US-10-170-235-28132	Sequence 28132, A
2	2324.5	89.8	1930	18	US-60-452-680-11396	Sequence 11396, A
3	2195.5	84.8	3294	15	US-10-170-235-28519	Sequence 28519, A
4	2195.5	84.8	3294	18	US-60-452-680-11397	Sequence 11397, A
5	2180.5	84.3	1389	12	PCT-US02-04915-41	Sequence 41, Appl
6	2180.5	84.3	1389	14	US-10-211-4652-41	Sequence 41, Appl
7	2180.5	84.3	1389	15	US-10-087-192-1137	Sequence 1127, Ap
8	2180.5	84.3	1389	14	US-10-126-052A-330	Sequence 330, App
9	2083.5	80.5	1297	15	US-10-087-192-1124	Sequence 1124, Ap
10	2017.5	78.0	2172	14	US-10-144-771-355	Sequence 355, App
11	1482	57.3	1359	2	PCT-US02-41414-1210	Sequence 1210, App
12	1482	57.3	1297	14	US-10-342-887-710	Sequence 710, App
13	1482	57.0	1359	2	PCT-US02-41414-1207	Sequence 1207, App
14	1474	57.0	1729	2	PCT-US02-41414-1206	Sequence 1206, App
15	1474	57.0	1729	14	US-10-144-771-10845	Sequence 10845, A
16	1468	56.7	2769	15	US-10-170-235-34391	Sequence 34391, A
17	1468	56.7	2769	15	US-09-949-016-4688	Sequence 4688, Ap
18	1465	56.6	2172	11	US-10-087-192-1126	Sequence 1126, Ap
19	809.5	31.3	212231	15	US-09-947-914-53	Sequence 53, Appl
20	809.5	31.0	6909021	9	US-10-087-192-1123	Sequence 1123, Ap
21	802	31.0	567	14	US-10-203-138A-6819	Sequence 6819, Ap
22	801	31.0	473	14	US-10-203-138A-8888	Sequence 8888, Ap
23	716	27.7	57726	11	US-09-949-016-16430	Sequence 16430, A
24	702.5	27.1	73725	2	PCT-US02-41414-1208	Sequence 1208, Ap
25	702.5	27.1	73725	2	PCT-US02-41414-1205	Sequence 1205, Ap
26	685.5	26.5	533	10	US-09-513-999C-1340	Sequence 1340, Ap
27	654	25.3	533	12	US-09-513-999C-1340	Sequence 1340, Ap
28	654	25.3	533	10	US-09-513-999C-1340	Sequence 1340, Ap
29	503	19.4	420	14	US-10-203-138A-3751	Sequence 3751, Ap
30	503	19.4	454	14	US-10-203-138A-1690	Sequence 1690, Ap
31	498	19.2	318	10	US-09-513-999C-25950	Sequence 25950, A
32	498	19.2	318	12	US-09-513-999C-25950	Sequence 25950, A
33	481	18.6	1752	15	US-10-219-051E-13871	Sequence 13871, A
34	474.5	18.3	1876	15	US-10-170-235-24888	Sequence 24888, A
35	474.5	18.3	1876	18	US-60-452-680-2295	Sequence 2295, App
36	474.5	18.3	1901	15	US-10-219-051E-13873	Sequence 13873, A
37	469	18.1	1581	15	US-10-170-235-34170	Sequence 34170, A
38	469	18.1	1581	19	US-60-455-444-673	Sequence 673, App
39	469	18.1	1581	19	US-60-465-241-673	Sequence 673, App
40	464	17.9	1165	15	US-10-170-235-24404	Sequence 24404, A
41	464	17.9	1165	18	US-60-452-680-296	Sequence 296, App
42	448	17.3	443	14	US-10-144-771-21735	Sequence 21735, A
43	444	17.2	2422	15	US-10-170-235-34449	Sequence 34449, A
44	444	17.2	2422	15	US-60-452-680-11304	Sequence 11304, A
45	443.5	17.1	3769	18	US-10-170-235-28517	Sequence 28517, A

## ALIGNMENTS

RESULT 1  
US-10-170-235-28132  
Sequence 28132, Application US/10170235  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig  
TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU.  
FILE REFERENCE: C1001380  
TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF  
CURRENT APPLICATION NUMBER: US/10/170, 235  
CURRENT FILING DATE: 2003-03-17  
NUMBER OF SEQ ID NOS: 42514  
SEQ ID NO 28132  
LENGTH: 1930  
TYPE: DNA  
ORGANISM: HUMAN  
US-10-170-235-28132

Alignment Scores:  
Pred. No.: 5.34e-259  
Score: 2324.50  
Percent Similarity: 89.36%  
Best Local Similarity: 87.55%

Length: 1930  
Matches: 436  
Conservative: 9  
Mismatches: 12

Query Match:	89.82%	Indels:	41
DB:	15	Gaps:	3

OY	21	AlaSerThrIleIleYsglnAlaLeuSerValIleSerGlnAspIleSerLeuPheGlnCys	21
Db	25	GCAGCTCATATACAGGAGACCCTTATACATTGTGACGAGAGACCAGTCGTTGTTGAGTGT	84
OY	22	AlaTyrGlySerProHisLeuAlaYsthrGluMetThrAlaSerSerSerGluTyr	41
Db	85	GCCTACGGAAACCCACACTGGCTTAAGACAGATGACCGCGCTCTCTCCAGCGACTAT	144
OY	42	GlyGlnThrSerLysMetSerProAlaGValProGlnGlnAspTrpLeuSerGlnProPro	61
Db	145	GGACGAGACTTCCAAAGTAGAGCCACCGCGTCCCTCACAGATGGTGGTCTCAACCCCA	204
OY	62	AlaArgValThrIleYsMetGluCysAsnProAsnGluValAsnGlySerArgAsnSer	81
Db	205	GGCAGGCTACCATCAATAATGGAAATGTAACTACCTAGCCAGGTGAATGGCTCAAGAACTCT	264
OY	82	ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly	101
Db	265	CCTGATGATATGACAGTGTGGCCAAAGCGGGAGATGTTGGCGACCCACACCGTGTGGG	324
OY	102	MetAsnTyrGlySerTyrMetGluGluYsthrHisIleProProAsnMetThrTrpAsn	121
Db	325	ATGAACCTACGACGCTACATGGAGAGAAAGCAATGCCACCCCAACATGACACAGAAC	384
OY	122	GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValAcGln	141
Db	385	GAGCCAGAGTTATGCTGTGCACACAGATCCTACGCTATGAGATACAGATCATGTGGCGAC	444
OY	142	TyrPheGluIuTrpAlaValLysGluTyrGlyLeuProAspValAlaPheLeuLeuPheGln	161
Db	445	TGGCTGGAGTGGCGGTGAATAATATGGCTCTCCAGAGTCAACATCTTGTATTTCACG	504
OY	162	AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro	181
Db	505	AACATCGATGGGAAGAACTGTGCCAAGATGACCAGAGACGACTTCCAGAGGCTCACCCCC	564
OY	182	SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluIuThrProLeuPro	201
Db	565	AGCTCAACAGCCGACATCCTTCTCTCACATCTCCATACCTCAGAGAGACTCCTCTTCCA	624
OY	202	HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla	221
Db	625	CATTGTACCTCGAGATGATGTATAAAGCTTACAAACATCTCCACAGGTTAAAGCATGCT	684
OY	222	ArgAsnThrGlyAlaIleThrPheIlePheProAsnThrSerValTyrProGluAlaThr	241
Db	685	AGAAACACAGGGGGTGCAGCTTTATTTTCCAAATACTCTCAGTATATCTCTGAAGCTACG	744
OY	242	GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaAlaArgSerAlaTrpThr	261
Db	745	CAAAAGAAATTACAACTAGCGCAGATTACCATATGAGCCCCCAGAGAGATCAGCTGAGCC	804
OY	262	SerHisSerHisProThr---GlnSerLysAlaIuGlnProSerSerSerThrTrpAlaPro	280
Db	805	GGTCAGGGCCACCCACGCGCCCAAGTCGAAAGTGTCTCAACCATCTCTTCCACAGTGGCC	864
OY	281	LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGluProThrSerSer	300
Db	865	AAACTGAAAGCACACGCTCCTCAGTTAGATCCTTATCAGATTCTTGACCAACAGTACG	924
OY	301	ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeu	320
Db	925	CGCCTTGAATATCCAGCGCAGTGGCCAGATCCAGCTTGGCAGATTCTCTCTGGAGGCTCTG	984
OY	321	SerAspSerSerAsnSerAsnCysIleIleThrTrpGluGlyThrAsnGlyLubPheLysMet	340
Db	985	TTGGACACCTTCAACTTCCAGCTGCATCCATGGGAAGCACCAAGGGAGTTTCAAGAAAG	1044

QY	341	thrasprroasbgluvalalalargargtrpblgluatyglusserysproasmetasn	360
Db	1045	ACGGATCCCGACGAGGTGGCCGGCGGGGAGAGCGGAAGACCAACCAATGTAAAC	1104
QY	361	tyrasplylseuseratargalaleuatytyrtyrtyrasplysasnilmetthrlysyal	380
Db	1105	TACGATAAGCTCAGCCGGCCCTCCGTTACTACTATGCAAGAACATCATACCAAGTTC	1164
QY	381	hisproproglusersemetyrlystyrproseraspleuoprotyrmetsetserlyr	400
Db	1164	-----	1164
QY	401	hisglylsargtyrtyralatyrltyrsphasaphehnsglylleaglnalaleuqlnpro	420
Db	1165	CATGGGAAGCGCTAGCGCTACAAAGTTGCACTTCCACGGGATCGCCAGGCCCTCCAGGCC	1224
QY	420	-----	420
Db	1225	CACCCCGGAGTCACTCTGTACAAAGTACCCTCAGACCTCCGTCATGGGCTCCAT	1284
QY	421	hisalahisproglinlysmetasnphavalalaprohisprofoalaleuproval1thr	440
Db	1285	CACGCCGCCACGAGAAATGAACTTGTGGCGCCGCCACCTCCAGCCCTCCCGTGACA	1344
QY	441	serserserphepaeialalaproasnprotyrtrpaaenserprothrglylyletry	460
Db	1345	TCTTCAGATTTTTTTTGGCGCCCAACCCATACACTGGAATTTACCAACGTGGGGGTATATAC	1404
QY	461	proasnthrargleuproalalalahismerproserhislseuglythryrtyr	478
Db	1405	CCCAACACTAGGCTCCGCCACGACGATGCTTCTCATCTGGGCACTTACTAC	1458

RESULT 2

US-60-452-680-11396

; Sequence 11396, Application US/60452680

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: GRUPE, Andrew

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01450

; CURRENT APPLICATION NUMBER: US/60/452,680

; CURRENT FILING DATE: 2003-03-07

; NUMBER OF SEQ ID NOS: 116213

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11396

; LENGTH: 1930

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-60-452-680-11396

Alignment Scores:

Pred. No.: 5,34e-259 Length: 1930

Score: 2324.50 Matches: 436

Percent Similarity: 89.36% Conservative: 9

Best Local Similarity: 87.55% Mismatches: 12

Query Match: 89.82% Indels: 41

DB: 18 Gaps: 3

US-09-902-772-4 (1-478) x US-60-452-680-11396 (1-1930)

QY 2 Alaserthrthrltylsglualalaleuservalvalserglunbspnserleupheglucys 21

Db 25 GCACGCTATATCAAGAGAACCTTATTCAGTTGTGAGTAGAGACCAAGTGTGTTGAAGT 84

QY 22 AAtyrglyserprohislseualalalythrglumethrthralaserSerSerSerglytyr 41

Db 85 GCCTACGAGAACGCCACACCTGGCTAGACAGAGATGACCGCGTCTCTCCAGAGACTAT 144

QY 42 Glylslnthrserlysmetserproalargvalproglinsasprtleuserglinpropro 61

Db 145 GGAGAGACTTCTCAAGATGAGAGCCACGCGTCTCTCAGCAGATGGGCTGTCAACCCCA 204



QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81  
 Db 205 GCCAGGGGACACATCAAAATGATGTAAACCTTAGCCAGGATGGATCGCTCAAGAACTCT 264  
 QY 82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGly 101  
 Db 265 CCTGATTAAGCACTGGGCGCAAGGCGGGAAGATGTGGGCGACCCAGACACCGTGGG 324  
 QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121  
 Db 325 ATGACATACGGCACTCATGAGAGAGAGACATGCAACCCCAACATGACCAAGAAC 384  
 QY 122 GluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArgGln 141  
 Db 385 GAGCGCAGAGTTATTCGTGCCAGCAGATCCTACGTATGAGATACAGACATGTCGGCAG 444  
 QY 142 TripleGluTrrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161  
 Db 445 TGGCTGGAGTGGGCGTGAAGAATATGCGCTTCCAAATGCAACATCTGTATATTCAG 504  
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181  
 Db 505 AACATTCATGGGAAGGAAGTGTGCAAGATGCAAGAGACGATTCACAGAGCTCACCCCC 564  
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisIleTyrLeuArgGluThrProLeuPro 201  
 Db 565 AGCTACAAACGGCGACATCTTCTCTCATCTCCACTACCTCAGAGAGACTCTCTTCCA 624  
 QY 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221  
 Db 625 CATTTGACTTCAGATGATGTGATTAAGCCTTACAAACCTCCACGCGTTAATGCATGCT 684  
 QY 222 ArgAspThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThr 241  
 Db 685 AGAAGACACAGGGGTGACGCTTTATTTTCCAAATCTTCAGATATCTCGAAGCTACG 744  
 QY 242 GlnArgIleThrThrArgProAspLeuProTyrGlnGlnIleAlaArgAspSerAlaTrrThr 261  
 Db 745 CAAGAATTTACAACTAGCCAGATTTACATATAGCCCCCGGAGAAATCAGCTTGAGCC 804  
 QY 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280  
 Db 805 GGTCCACGGCCACCCACGCCAGTCGAAAGCTCTCAACCATCTCTCCACAGTGCCTC 864  
 QY 281 LysThrGlnLysGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300  
 Db 865 AAAACTGAAGACAGCGCTCTCACTAGATCTTATCAGATCTTGAGCAACAAGATGAGC 924  
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrrPglPheLeuLeuGluLeuLeu 320  
 Db 925 CGCCTTGCAAAATCCAGGACAGTGGCCAGATCCAGCTKTGGCAGTTCCCTCGGAGCTCTG 984  
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTrrPglLysThrAsnGlyLysPheLysMet 340  
 Db 985 TCGGACAGCTCCAACTCAGCTGATCAGCTGGGAAGCAACAAGGGGAGTTCAAGATG 1044  
 QY 341 ThrAspProAspGluValAlaArgArgTrrPglLysGlnArgLysSerLysProAsnMetAsn 360  
 Db 1045 ACGGATCCGACGAGGTGGCGCGCTGGGAGAGGGAAGCAAAACCAACATTAATGAC 1104  
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTrrTyrTrrAspLysAsnIleMetThrLysVal 380  
 Db 1105 TAGCATTAAGCTCAGCGCGCTCCGTTACTACTATACAAAGAATCATGACCAAGGTC 1164  
 QY 381 HisProProGluSerSerMetTrrLysTrrProSerAspLeuProTrrMetSerSerTyr 400  
 Db 1164 ----- 1164  
 QY 401 HisGlyLysArgTrrAlaTrrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420  
 Db 1165 CATGGAGGCGCTTAGCGCTTACAGTTCGACTTCACGGGATTCGCCAGGCGCTTCACGCC 1224  
 QY 420 ----- 420

Db 1225 CACCCCGGAGTCATCTCTGACAAAGTACCCCTCAGACCTCCCGTACATGGGCTCCAT 1284  
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440  
 Db 1285 CACGCCACCCACGGAAGATGATCTTGTGGCGGCCACCTCAGCGCTCCCGTGACA 1344  
 QY 441 SerSerSerPhePheAlaAlaProAsnProTrrTrrAsnSerProThrGlyLysIleTyr 460  
 Db 1345 TCTTCCAGTTTCTTGTGCTGCCCAAAACCATCTGGAATTCACCACTGGGGGATATATC 1404  
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTrrTyr 478  
 Db 1405 CCCAACAAGTGGCTCCCGACAGCCATATGGCTTCTCATCTGGGCGACTTACTAC 1458  
 RESULT 3  
 US-10-170-235-28519  
 : Sequence 28519, Application US/10170235  
 : GENERAL INFORMATION:  
 : APPLICANT: VENTER, J. Craig  
 : TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU  
 : : FILE REFERENCE: CL001380  
 : : CURRENT FILING DATE: 2003-03-17  
 : : NUMBER OF SEQ. ID NOS: 42514  
 : SEQ ID NO 28519  
 : LENGTH: 3294  
 : TYPE: DNA  
 : ORGANISM: HUMAN  
 US-10-170-235-28519  
 Alignment Scores:  
 Pred. No.: 1,48e-243 Length: 3294  
 Score: 2195.50 Matches: 416  
 Percent Similarity: 84.97% Conservative: 8  
 Best Local Similarity: 83.37% Mismatches: 10  
 Query Match: 84.83% Indels: 65  
 DB: 15 Gaps: 4  
 US-09-902-772-4 (1-478) x US-10-170-235-28519 (1-3294)  
 QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 94 ATGGCCAGCACTATTAAAGAACTTATCACTGTGTGATGAGACCACTGTTGTTGAG 153  
 QY 21 CysAlaTrrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 Db 154 TGTGCTTACGGAAGCCACACCTGGCTTAAGACAGATGACCGGCTCCTCCACGAGAC 213  
 QY 41 TrrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrrPleuSerGlnPro 60  
 Db 214 TATGGACAGACTTCAAGATGAGCCACGGCTCCCTCAGAGGATTTGGCTTCAACCC 273  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 274 CCAAGCCAGGTCACATCAAAATGGAATGTAAACCTTAGCCAGTGAATGGCTCAAGAAC 333  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 Db 334 TCTCCTGATGAATGACACTGTGGCCAAAGGCGGAAGATGTGGGACACCCAGACACCGTT 393  
 QY 101 GlyMetAsnTrrGlySerTrrMetGluGluLysHisIleProProProAsnMetThrThr 120  
 Db 394 GGGATGAACATACGCACTCATGAGAGAACACATGTCACCCCAACATGACACAGC 453  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrrSerThrAspHisValArg 140  
 Db 454 AACGAGGCGCAGATTAATTCGTGCCAGCAGATCCTACGTATGAGATACAGACCATGTGCGG 513  
 QY 141 GlnTrrPleuGluTrrPalaValLysGluTrrGlyLeuProAspValAspIleLeuPhe 160  
 Db 514 CAGTGGCTGGAGTGGGCGTGAAGAATATGCGCTTCCAGACGTCACATCTGTATATTC 573

QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 DB 574 CAGAAATCGATGGAGGAGAACTGTGCNAGATGACCAAGACGACTTCACAGGCTCAC 633  
 QY 181 ProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisLysThrLeuArgGluThrProLeu 200  
 DB 634 CCCAGCTACAAACGCGGACATCTCTCTCACAATCTCCACTACGAGAGACATCTCTT 693  
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220  
 DB 694 CCACATTGACTTCAGATGATGTGATTAACCTTACAAACTCTCCACGGTTAATGAT 753  
 QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240  
 DB 754 GCTAGAAACACA----- 765  
 QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyr 260  
 DB 766 -----GATTACCAATGAGCCCCAGAGAGATCAGCCTGG 801  
 QY 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThrVal 279  
 DB 802 ACCGGTACAGCCGACCCACGCCCCAGTGAAGCTGCTCAACCATCTCTTCCACAGTG 861  
 QY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299  
 DB 862 CCCAAACTGAAAGCACGCGCTCTCAGATTAGATCTTATCAGATTCTTGACCAACAAGT 921  
 QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuLeuGluLeu 319  
 DB 922 AGCGGCTTGCAAAATCCAGCGAGTGCGCAGATCCAGCTTGGCAGTTCCTCTGGAGTTC 981  
 QY 320 LeuSerAspSerSerAsnSerAsnLysIleThrTrpGluGlyThrAsnGlyLysPheLys 339  
 DB 982 CTGTCGGACACTCCCACTCAGTGCATCAGCTGGAGGACACCAACGGGAGTTTCAG 1041  
 QY 340 MetThrAspProAspGluValAlaArgTrpGlyGluArgLysSerLysProAsnMet 359  
 DB 1042 ATGACGAGATCCCGAGGAGTGCGCGCGCGGAGAGACGGAAGACCAACCAACATG 1101  
 QY 360 AsnThrAspLysLeuSerArgAlaLeuArgTyrTyrTrpAspLysAsnIleMetThrLys 379  
 DB 1102 AACACAGATAGGCTACGCGCGCTCGTACTACTGATGACAAAGAACATCATACCAAG 1161  
 QY 380 ValHisProProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSer 399  
 DB 1162 GTC----- 1164  
 QY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419  
 DB 1165 ---CATGGAAAGCGCTACGCTACAGTTCGACTTCACAGGGATGCGCCAGGCTTCAG 1221  
 QY 420 Pro----- 420  
 DB 1222 CCCCACCCCGGAGTATCTGTACAACTACCCCTCAGACTCCCGTACAGGGCTCC 1281  
 QY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439  
 DB 1282 TATCAGGCCCAACAGAGAGATGAACTTGTGGCGGCCCACTCCAGGCTTCAGGCTG 1341  
 QY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLys 459  
 DB 1342 ACAATCTTCAGATTTTGTGCTGCCCAACCACTACGGAATTCACCAACTGGGGGTATA 1401  
 QY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478  
 DB 1402 TACCCCAACACTAGGCTCCACAGCATATAGCTTCTCATCTGGGCACTTACTAC 1458  
 RESULT 4  
 US-60-452-680-11397  
 ; Sequence 11397, Application US/60452680  
 ; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele  
 ; APPLICANT: GROPE, Andrew  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C1001450  
 ; CURRENT APPLICATION NUMBER: US/60/452,680  
 ; NUMBER OF INVENTION: 2003-03-07  
 ; NUMBER OF SEQ ID NOS: 116213  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 11397  
 ; LENGTH: 3294  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-60-452-680-11397  
 Alignment Scores:  
 Pred. No.: 1,48e-243 Length: 3294  
 Score: 2195.50 Matches: 416  
 Percent Similarity: 84.97% Conservative: 8  
 Best Local Similarity: 83.37% Mismatches: 10  
 Query Match: 84.83% Indels: 65  
 DB: 18 Gaps: 4  
 US-09-902-772-4 (1-478) x US-60-452-680-11397 (1-3294)  
 QY 1 MetaLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 DB 94 ATGCCAGCAGCACTATTAGAAAGCCTTATCAGTTGTGATGAGGACCAAGTGGTTGAG 153  
 QY 21 CysAlaIleArgLysSerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 DB 154 TGTGCTACAGGAAGCCACACCTGGCTTAAGACAGAGATGACCGGCTCTCTCCAGGAC 213  
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnAlaSerTrpLeuSerGlnPro 60  
 DB 214 TATGACAGACTTCCAGAGTAGAGGCCAGCGCTCCCTAGCAGATGGCTGTCTCAACCC 273  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 DB 274 CCAGCCAGGCTCACACATCAAAATGGAATGTAACCTTACCCAGGAGTAAGTGCACAGGAC 333  
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAspVal 100  
 DB 334 TCTCTGATGATATGCACTGTGGCCAAAGGGGGAAGTGGTGGGACACACACCGTT 393  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluCysHisIleProProProAsnMetThrThr 120  
 DB 394 GGGATGAACTACGGCAGCTACATGAGGAGAAAGCACATGCCACCCCAACATGACCAG 453  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrPheThrAspHisValArg 140  
 DB 454 AACGAGGCGAGAGTATCGTGCCAGCAGATCTTACGCTATGAGATGACAGACATGTGCGG 513  
 QY 141 GlnThrLeuGlnThrPheAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160  
 DB 514 CAGTGGCTGGAAGTGGGGGTGAAGAATATGAGGCTTCCAAATGTCAACATCTGTATTC 573  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 DB 574 CAGAAATCGATGGAGGAGAACTGTGCNAGATGACCAAGACGACTTCACAGGCTCAC 633  
 QY 181 ProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisLysThrLeuArgGluThrProLeu 200  
 DB 634 CCCAGCTACAAACGCGGACATCTCTCTCACAATCTCCACTACGAGAGACATCTCTT 693  
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220  
 DB 694 CCACATTGACTTCAGATGATGTGATTAACCTTACAAACTCTCCACGGTTAATGAT 753  
 QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240  
 DB 754 GCTAGAAACACA----- 765

```

Oy 241 ThrGlnArgIleThrThrArgProAspLeuProTyGlnGlnAlaArgSerAlaTrp 260
Db 766 -----GATTACCATATGAGCCCCAGAGATTCAGCCGG 801
Oy 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThrVal 279
Db 802 ACCGGTCAAGCCAGCCAGCCAGCCAGCGAAGCTGCTCAACATCTCTTCCACAGTG 861
Oy 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyGlnIleLeuGlyProThrSer 299
Db 862 CCCAAACTGAAGACCGCTCTCAGTACATCTTATGATTCCTTATGAGTTCGACCAACAGT 921
Oy 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPglPheLeuLeuGlnLeu 319
Db 922 AGCCGCTTGCAGATCCAGGAGTGGCCAGATCCAGCTKRTGGCAGTTCTCTGGAGCTC 981
Oy 320 LeuSerAspSerSerAsnSerAsnLysIleThrTrpGluGlyThrAsnGlyLupheLys 339
Db 982 CTGTGCGACAGCTCCACTCCAGCTGCATCAGCTGGGAGGACCAACGGGAGTTCAG 1041
Oy 340 MethrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 359
Db 1042 ATGAGGATCCCGAGAGGTGGCCGGCGTGGGAGAGCGAAGAGCAACCAACATG 1101
Oy 360 AsnTyraAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
Db 1102 AACACAGTATAGTCAGCGCGGCGCTCGTACTACATGACAAAGAACATATGACCAAG 1161
Oy 380 ValHisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399
Db 1162 GTC----- 1164
Oy 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
Db 1165 ---CATGGAGAGCGCTACGCTACAGTTCACATTCACGCGGATGCGCCAGGCCCTCAG 1221
Oy 420 Pro----- 420
Db 1222 CCCACCCCGCGAGTCACTCTGTACAACTACCCCTCAGACCTCCGTACATGGGCTCC 1281
Oy 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439
Db 1282 TATACAGCCCAACCAACAAATGAACTTGTGGCGGCCCACTCCAGCCCTCCCGCTG 1341
Oy 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyLe 459
Db 1342 ACATCTTCCACTTTTTCCTGCTGCCCCCAACCCATACCTGGAATTCACCAACTGGGGGTATA 1401
Oy 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1402 TACCCCAACTAGGCTCCCAACAGCATATGCTCTCTGCTGGGCACTTACTAC 1458

```

RESULT 5  
PCT-US02-04915-41  
Sequence 41, Application PC/TUS0204915

GENERAL INFORMATION:  
APPLICANT: Murray, Richard  
APPLICANT: Glynn, Richard  
APPLICANT: Watson, Susan R.

APPLICANT: Aziz, Natasha  
APPLICANT: Eos Biotechnology, Inc.

TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and  
TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators

FILE REFERENCE: 018501-006200PC  
CURRENT APPLICATION NUMBER: PCT/US02/04915

CURRENT FILING DATE: 2002-02-14  
PRIOR APPLICATION NUMBER: US 09/7784, 356

PRIOR FILING DATE: 2001-02-14  
PRIOR APPLICATION NUMBER: US 09/791, 390

PRIOR FILING DATE: 2001-02-22  
PRIOR APPLICATION NUMBER: US 60/285, 475

PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: US 60/310, 025

```

; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/350, 666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334, 244
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-04915-41

Alignment Scores:
Pred. No.: 1,75e-242 Length: 1389
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: Gaps: 4

US-09-902-772-4 (1-478) x PCT-US02-04915-41 (1-1389)

Oy 2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluLys 21
Db 25 GCAGCTCATATCAAGAGAGCCCTTACATTTGATGAGACAGCAGCTCTTTGAGCTGT 84
Oy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlyTyr 41
Db 85 GCTTACGAGAACGCCACCTGCTGTAAGACAGATGACCCGCTCTCTCCAGGAGTAT 144
Oy 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
Db 145 GGACAGACTTCCAGATGAGGCCACGCTCCCTCAGCAGGATGGCTGTCTCAACCCCA 204
Oy 62 AlaArgValThrIleLysMetGluLysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 205 GCCAGAGGTACACATCAAAATGAAATGTAACCTGACGCGATGAGTATGGCTCAAGAACTCT 264
Oy 82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGly 101
Db 265 CCTGATGAAATGCACTGTGGCCAAAGCGGGAAGATGGGGGACCCCAACACCGGTGG 324
Oy 102 MetAsnTyrGlySerTyrMetGluLysHisIleProProProAsnMetThrAsn 121
Db 325 ATGAACTACGCGACCTACATGAGAGAGAGACATGACACCCCAACATGACACGAGAC 384
Oy 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db 385 GAGCGCAGAGTTATCGTCCAGCAGATCCCTACGCTATGAGTACAGACATGTCGGGAG 444
Oy 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161
Db 445 TGGCTGAGAGTGGGGGGAAGAAATATGGCCCTTCAGACGTCACATCTTGTATTCCAG 504
Oy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 505 AACATTCATGGGAAGAGACTGTGCAAGATGACCAAGACACTTCCAGAGGCTCAACCCC 564
Oy 182 SerTyraAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluTrpProLeuPro 201
Db 565 AGTACAAAGCGCGACATCTCTCTCCATCTCCACATGACAGAGAGACTCTCTTCA 624
Oy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db 625 CATTTGACTTCAGATGATGTTGATTAAGCCCTTACAAAACCTCCACGGTTAATGCAATGCT 684
Oy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db 685 AGAAACACA----- 693
Oy 242 GlnArgIleThrThrArgProAspLeuProTyGlnGlnAlaArgSerAlaTrpThr 261

```

```

Db 694 -----GATTACCATATGAGCCCCCAGAGATCAGCCTGAGCC 732
Oy 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 723 GGTACAGGCGCAACCCAGCCCGCCAGTGAAGCTGTCTCAACATCTCTCCACATGCGCC 792
Oy 281 LysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db 793 AAAACTGAAGACAGCGCTCTCAGTTAGATCCTTATCAGATTCCTTGAGCAACAAGTAGC 852
Oy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPheIleLeuGlnLeuLeu 320
Db 853 CGCCTTGCATAATCCAGCGCATGCCAGATCCAGCTTGCGAGTCTCTCTGAGCTCTG 912
Oy 321 SerAspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyIlePheLysMet 340
Db 913 TCGGACAGCTCCACATCCAGCTGATCAGTCCGGAAGGACACAGGGGAGTTCAGATG 972
Oy 341 ThrAspProAspGlnValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsn 360
Db 973 ACGGATCCCGACGAGGTGGCCGCGCTGGGAGAGCGGAAGAGCAACCCACATGAGAC 1032
Oy 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db 1033 TACGATTAAGCTCAGCGCGCCCTCCGTTACTACTACTATGACAGAACATCATGACCAAGTTC 1092
Oy 381 HisProGlnUserSerMetLysTyrTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1092 ----- 1092
Oy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1093 CATGGAGAGCCCTACGCTACAGTTCAGCTTCCACAGGGATGCGCCAGGCCCTCCAGGCC 1152
Oy 420 ----- 420
Db 1153 CACCCCCGGAGTACTCTGTACAGATCACCCCTCAGACCTCCGCTACATGGGCTCTAT 1212
Oy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1213 CACGGCCACCCAGAAAGATTAATTTGTGGCGCCGCCCTCCAGCCCTCCCGTGACA 1272
Oy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysTyr 460
Db 1273 TCTTCCAGTTTTTTTGTGTGCCCCCAACCATCTGGAATTCACCAACCTGGGGTATATAC 1332
Oy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1333 CCCAACACTAGGCTCCCGACAGCATATGCTTCTCATCTGGGCACTTACTAC 1386

RESULT 6
: Sequence 41, Application US/10211462
: GENERAL INFORMATION:
: APPLICANT: Murray, Richard
: APPLICANT: Glynn, Richard
: APPLICANT: Watson, Susan R.
: APPLICANT: Aziz, Natasha
: APPLICANT: Eos Biotechnology, Inc.
: TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
: TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
: FILE REFERENCE: 018501-006200US
: CURRENT FILING DATE: US/10/211,462
: PRIOR APPLICATION NUMBER: US 09/784,356
: PRIOR FILING DATE: 2001-02-14
: PRIOR APPLICATION NUMBER: US 09/791,390
: PRIOR FILING DATE: 2001-02-22
: PRIOR APPLICATION NUMBER: US 60/310,025
: PRIOR FILING DATE: 2001-08-03
: PRIOR APPLICATION NUMBER: US 60/334,244
: PRIOR FILING DATE: 2001-11-29
: NUMBER OF SEQ ID NOS: 230

```

```

: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 41
: LENGTH: 1389
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-211-462-41

Alignment Scores:
Score: 1.75e-242 Length: 1389
Pred. No.: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 14 Gaps: 4

US-09-902-772-4 (1-478) x US-10-211-462-41 (1-1389)

Oy 2 AlaSerThrIleLysGlnAlaLeuSerValSerGlnAspGlnSerLeuPheGlnCys 21
Db 25 GCAGCTCATATCAAGAGGAGCCCTTATCAGTTGTGAGTGAGCCAGTGTGTTGTGAGTGT 84
Oy 22 AlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerSerGlyTyr 41
Db 85 GCCTACGGAAGCCACACCTGGTAAAGACAGATGACCGGCTCTCTCCAGCGACTAT 144
Oy 42 GlynThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
Db 145 GGACAGACTTCCAGAGTACAGCCAGCGGCTCCAGCAGATGCGTGTCTCAACCCCA 204
Oy 62 AlaArgValThrIleLysMetGlnCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 205 GCCAGGGGTCACACCAAAATGGAATGTAACTTACCCAGGTGAATGGCTCAAGGAACTCT 264
Oy 82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGly 101
Db 265 CTTGATGATGACATGTGGCCAAAGCGGGAAGTGGTGGCCAGCCAGACCCGTTGGG 324
Oy 102 MetAsnTyrGlySerTyrMetGlnGlnLysHisIleProProProAsnMetThrThrAsn 121
Db 325 ATGAACATACGGCAGCTCATGAGAGAGACAGACATGCGACCCCAAAACATGACACAGAAC 384
Oy 122 GlnArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db 385 GAGCGAGAGTATTCGTCGCCAGCATCTTACGCTATGAGTACAGACATGTGGCGCAG 444
Oy 142 TrpLeuGlnTrpAlaValLysGlnTyrGlyLeuProAspValAspIleLeuPheGln 161
Db 445 TGGCTGGAGTGGGGGTAAGAAATATAGGCTTCCAGACGTCACATCTTGTATTTCAG 504
Oy 162 AsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 505 AACATCGATGGAGAGAACTGTGCAGATGACCAAGAGACGCTCCAGAGGCTCACCC 564
Oy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGlnThrProLeuPro 201
Db 565 AGCTTACAAAGCGGCATCTTCTTCCATCTCCACATCTCCAGAGAGACTCTCTTCCA 624
Oy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db 625 CATTGACTTCAGATGATGTGTAAGAACCTTACAAACCTTCACAGCGTTAATGCAATGCT 684
Oy 222 ArgAsnThrGlyGlnAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThr 241
Db 685 AGAAGACACA----- 693
Oy 242 GlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrpThr 261
Db 694 -----GATTACCATATGAGCCCCCAGAGATCAGCCTGAGCC 732
Oy 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 723 GGTACAGGCGCAACCCAGCCCGCCAGTGAAGCTGTCTCAACATCTCTCCACATGCGCC 792

```

```

Qy 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleuGlyProThrSerSer 300
Db 793 AAAACTGAAGACCGGCTCCCTACAGTACATCTTATCGATTCTTGGACCAAGATGAC 852
Qy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuGlnLeuLeu 320
Db 853 CGCCTTGCAATTCAGGAGCAGTGGCCAGATCCAGCTTTGGAGTTCCCTCCGAGCTCTG 912
Qy 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphLeuMet 340
Db 913 TCGGACAGCTCCAACTCCAGTCGATCACCGGAGAGCCAAACGGGAGTTCAAGATG 972
Qy 341 ThrAspProAspGluValAlaArgArgTyrGlyGlnArgGlySerLysProAsnMetAsn 360
Db 973 ACGGATCCCGACGAGAGTGGCCCGGCTGGGGAGACGGAGAGCAAAACCAACATGAC 1032
Qy 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db 1033 TACGATTAAGTCACACCGCGCCCTCGTACTACTATGACAGAACATCATCATCAGCAG 1092
Qy 381 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1092 ----- 1092
Qy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1093 CATGGAGAGCGCTACGCTTACAGTTCGACTTCCACGGGATGCCCGCCCTCAGGCC 1152
Qy 420 ----- 420
Db 1153 CACCCCGGAGTATCTCTGTACAGTACCCCTCAGACTCCGCTACATGGGCTCTAT 1212
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1213 CACGCCCAACCCAGAGATGAATCTGTGGCCCAACCTCCAGCCCTCCCGTGACA 1272
Qy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 460
Db 1273 TCTTCCAGTTTCTTGGTGGCCCAACCCATCTGGAATTCACCAACGGGGGATATATAC 1332
Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1333 CCCAACACTAGGCTCCGCCACAGCATATGCTTCTCATCTGGGACACTTACTAC 1386

RESULT 7
US-10-087-192-1127
: Sequence 1127, Application US/10087192
: GENERAL INFORMATION:
: APPLICANT: Morris, David W.
: APPLICANT: Engelhard, Eric K.
: TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: CANCER
: FILE REFERENCE: 52945200122
: CURRENT APPLICATION NUMBER: US/10/087,192
: PRIOR FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 09/747,377
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/798,586
: PRIOR FILING DATE: 2001-03-02
: NUMBER OF SEQ ID NOS: 2059
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1127
: LENGTH: 1389
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-087-192-1127

Alignment Scores:
Pred. No.: 1,75e-242 Length: 1389
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65

```

```

DB: 15 Gaps: 4
US-09-902-772-4 (1-478) x US-10-087-192-1127 (1-1389)
Qy 2 AlaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerIleuPheGluCys 21
Db 25 GCAGCTCATATCAAGAGAGGAGCTTATCATCTTGATGAGTGAAGACCAAGTCTGTTGAGAGCT 84
Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGluTyr 41
Db 85 GCTTACGAGAACGCCACACCTGCTTAAGACAGATGACAGCGGCTCTCTCCAGGACATAT 144
Qy 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
Db 145 GGACAGACTTCAAGATGATGAGCCAGCTCCCTCAGGAGATGGCTGTCTCAACCCCA 204
Qy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 205 GCCAGGGTCACATCAATAAATGGAATGTAACTTACGCCAGTGGAATGGCTCAAGAACTCT 264
Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValIly 101
Db 265 CCTGATGATGACAGTGTGGCCAAAGCCGGAGAGATGTGTGGCAGCCAGACACCGTTGG 324
Qy 102 MetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrAsn 121
Db 325 ATGAACTAGGAGCTTACATGAGAGAGACACATGCCACCCCAACATACACACGAAAC 384
Qy 122 GluArgArgValIleValProAlaAspProThrIleuTyrSerThrAspHisValArgGln 141
Db 385 GAGCGAGAGATTATCGTCCAGACAGATCTCAGCTATGATGATAGACATGAGCATGTGGGAG 444
Qy 142 TrpLeuGlnTyrPalaValLysGlyLysTyrGlyLeuProAspValAspIleLeuPheGln 161
Db 445 TGGCTGGAGTGGGGGTAAAGAAATATGCTTCCAGACGTCACATCTTGTATTCAG 504
Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 505 AACATCGATGGAGAGAACTGTGCAAGATGACCAAGACGACTTCCAGAGGCTCACCC 564
Qy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGlyLuphProLeuPro 201
Db 565 AGTAAACAAGCGGACATCTCTCTCATCATCTCCACTCTCAGAGAGACTCTCTTCCA 624
Qy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db 625 CATTTGACTTCAAGATGATGTGATTAACCTTACAAACTCCACAGGCTTAATCATCTCT 684
Qy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db 685 AGAAACACA----- 693
Qy 242 GlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTyrThr 261
Db 694 -----GATTTACCATATGAGCCCGCCAGAGAGATGAGCTTGAGACC 732
Qy 262 SerHisSerHisProThr--GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 733 GGTACAGGCGCACCCACAGCCCGCCAGTCGAAACCTCTCAACATCTCTCCACAGATGCC 792
Qy 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleuGlyProThrSerSer 300
Db 793 AAAACTGAAGACCGGCTCCCTACAGTACATCTTATCGATTCTTGGACCAAGATGAC 852
Qy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuGlnLeuLeu 320
Db 853 CGCCTTGCAATTCAGGAGCAGTGGCCAGATCCAGCTTTGGAGTTCCCTCCGAGCTCTG 912
Qy 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphLeuMet 340
Db 913 TCGGACAGCTCCAACTCCAGTCGATCACCGGAGAGCCAAACGGGAGTTCAAGATG 972
Qy 341 ThrAspProAspGluValAlaArgArgTyrGlyGlnArgGlySerLysProAsnMetAsn 360

```

```

Db      973  ACGGATCCCGACGAGTGCGCCGCGTGGGAGAGCGAAGCAACCAACATGAC 1032
Qy      361  TTTAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db      1033  TACATAGAGCTCACCGCGCCCTCCGTTACTACTAGTCAAGAACATCATACCAAGGTC 1092
Qy      381  HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db      1092  ----- 1092
Qy      401  HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1093  CATGGAGACGCTACGCGCTTACAGATTCCACGAGGATGCCAGGCCCTCCAGGCC 1152
Qy      420  ----- 420
Db      1153  CACCCCGGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGCTCCAT 1212
Qy      421  HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db      1213  CACGCCACCCACAGAGATGACTTGTGGCGCCGCCACCTCCAGGCCCTCCCGTGACA 1272
Qy      441  SerSerSerPhePheAlaAlaProAsnProTyrTyrPheSerProThrGlyIleTyr 460
Db      1273  TCTTCAGTTTTTTGTGCTGCCCAACCCATACGTGAATTCCACCACTGGGGGTATATAC 1332
Qy      461  ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleTyr 478
Db      1333  CCCAACACTAGCGTCCCGCCACGACCATATGCTTCTCATCTGGGCACTTACTAC 1386

```

## RESULT 8

```

US-10-126-052A-330
: Sequence 330, Application US/10126052A
: GENERAL INFORMATION:

```

```

: APPLICANT: Mutray, Richard
: APPLICANT: Eos Biotechnology, Inc.
: TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
: FILE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
: FILE REFERENCE: 018501-001530US
: CURRENT APPLICATION NUMBER: US/10/126, 052A
: PRIOR FILING DATE: 2002-04-18
: PRIOR APPLICATION NUMBER: US 60/284, 770
: PRIOR FILING DATE: 2001-04-18
: PRIOR APPLICATION NUMBER: US 60/290, 492
: PRIOR FILING DATE: 2001-05-10
: PRIOR APPLICATION NUMBER: US 60/339, 245
: PRIOR FILING DATE: 2001-11-09
: PRIOR APPLICATION NUMBER: US 60/350, 666
: PRIOR FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/334, 370
: PRIOR FILING DATE: 2001-11-29
: PRIOR APPLICATION NUMBER: US 60/372, 246
: PRIOR FILING DATE: 2002-04-12
: NUMBER OF SEQ ID NOS: 691
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 330
: LENGTH: 3166
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-126-052A-330

```

## Alignment Scores:

```

Pred. No.: 7.69e-242 Length: 3166
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 14 Gaps: 4

```

```

US-09-902-772-4 (1-478) x US-10-126-052A-330 (1-3166)

```

```

Qy      2  AlaSerThrIleLysGlnAlaLeuSerValValSerGlnAspGlnSerIlePheGluCys 21
Db      281  GCAGCTCATATCAAGGAAGCTTATCAGTTGTGATGAGACAGCATGTTGTTGAGAGCT 340
Qy      22  AlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGluTyr 41
Db      341  GCCTACGAGAACGCCACCTGGCTTAACAGAGATGACCGGCTCTCTCCAGGACTAT 400
Qy      42  GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 61
Db      401  GGACAGACTTCCAGATGAGCCACCGGCTCCCTCAGACGATTTGGCTGTCTCAACCCCA 460
Qy      62  AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db      461  GCCAGGGTACCATCAAAATGATGATTAACCTTACCCAGGATGATGCTCAAGAACTCT 520
Qy      82  ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGly 101
Db      521  CCTGATGATGCAATGCTGTGGCCAAAGCGGGAAGATGGTGGGACCCAGACCGCTTGG 580
Qy      102  MetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThrAsn 121
Db      581  ATGAACTACGGCAGCTTACATGAGAGACAGACATGCCACCCCAAAACATGACACGAC 640
Qy      122  GlnArgArgValIleValProAlaAspProThrIleuTyrPheThrAspHisValArgGln 141
Db      641  GAGCGCAGAGTTATCGTCCGACGAGATCTCTACCTATGAGATGACAGACATGTGCGG 700
Qy      142  TrpLeuGluTyrPheValValLysGluTyrGlyLysProAspValAlaPheLeuPheGln 161
Db      701  TGGCTGAGATGGGGGGTGAAGATATGCGCTTCCACACGCAACATCTGTATTTCAG 760
Qy      162  AsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db      761  AACATCATGGAGAGAACTGTGCAACATGACCAAGACAGACTCCAGAGAGCTACACCCC 820
Qy      182  SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro 201
Db      821  AGCTACAAAGCGCGACCTTCTCTCACAATCTCCACATCTCCAGAGAGACTCTCTTCCA 880
Qy      202  HisLeuThrSerAspAspValAspLysValLeuGlnAsnSerProArgLeuMetHisAla 221
Db      881  CATTTGACTTCAGTATGATGTGATTAAGCCCTTACAAAACCTCTCCAGCGTTAATGAC 940
Qy      222  ArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThr 241
Db      941  AGAAGACACA----- 949
Qy      242  GlnArgIleThrThrArgProAspLeuProTyrGlnAlaArgAlaArgSerAlaTyrThr 261
Db      950  -----GATTTACCATATGACGCCCCCAAGAGATGACCTTGAGCC 988
Qy      262  SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db      989  GGTTCAGCGCCACCCACAGCCCAAGTCAAAAGCTCTCAACATCTCTTCCACAGTGGCC 1048
Qy      281  LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db      1049  AAAACTGAAGACACGCGTCTCTCACTTATGATCATCTTATGATGATGATGATGATGAT 1108
Qy      301  ArgLeuAlaAsnProGlnSerGlyGlnIleGlnLeuThrPheLeuLeuGlnLeuLeu 320
Db      1109  CGCCTTGCAAAATCGAGCACTGGCCACAGATCCACCTTTGGAGATTCCTCCGAGCTCT 1168
Qy      321  SerAspSerSerAsnSerAsnCysIleThrTyrGlnGluGlyHisAsnGlyLysMet 340
Db      1169  TCGGACAGCTTCCAACTCCACCTGATCCTGGGAAGGCAACCAAGGGAGTTCAAGATG 1228
Qy      341  ThrAspProAspGlnValAlaArgArgTyrGlyLysLysSerLysProAsnMetAsn 360
Db      1229  ACGGATCCGACAGCGAGTGGCGCGCTGGGAGAGCGAAGCAACCAACCAATGAAAC 1288
Qy      361  TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380

```

```

|||||
Db 1289 TACATAGCTACACCCGCGCTCCGCTACTACTATGACAAGACATCATGACCAAGTCC 1348
QY 381 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1348 -----
QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1349 CATGGGAAGCGCTACGCGCTACAACTTGACTTCACAGGAGATCGCCAGCGCTCCACGCC 1408
QY 420 -----
Db 1409 CACCCCGGAGTCATCTGTACAGTACCCCTCAGACCTCCCGTACATAGGCTCTCAT 1468
QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1469 CACGCGCACCCAGAAAGATGAACTTGTGGCGCCACCCCTCCAGCCCTCCCGTACA 1528
QY 441 SerSerPhePheAlaAlaProAsnProTyrTyrPheSerProThrGlyIleTyr 460
Db 1529 TCTTCAGTTTGTGTGCTGCCCAACCCATCTGGAATTCACCACTGGGGGTATATAC 1588
QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1589 CCCAACACTAGGCTGCCCGCCACGACCATATGCTTCATCTGCGACTACTAC 1642

RESULT 9
US-10-087-192-1124
: Sequence 1124, Application US/10087192
: GENERAL INFORMATION:
: APPLICANT: Morris, David W.
: APPLICANT: Engelhard, Eric K.
: TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: CANCER
: FILE REFERENCE: 529452000122
: CURRENT APPLICATION NUMBER: US/10/087,192
: CURRENT FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 2000-12-22
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/798,586
: PRIOR FILING DATE: 2001-03-02
: NUMBER OF SEQ ID NOS: 2059
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1124
: LENGTH: 1297
: TYPE: DNA
: ORGANISM: Mus musculus
US-10-087-192-1124

Alignment Scores:
Pred. No.: 3,07e-231 Length: 1297
Score: 2083.50 Matches: 394
Percent Similarity: 82.11% Conservativeness: 10
Best Local Similarity: 80.08% Mismatches: 7
Query Match: 80.51% Indels: 81
Gaps: 3

US-09-902-772-4 (1-478) x US-10-087-192-1124 (1-1297)
QY 7 GluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCysAlaTyrGlySerPro 26
Db 2 GAGCGCTTGTACAGTTGTGAGCGAGACGACGACTATTTGAGTGTGCTTACGGAACGCA 61
QY 27 HisLeuAlaLysThrGlnMetThrAlaSerSerSerSerGlnTyrGlyGlnThrSerLys 46
Db 62 CACCTGGCTAAGAGAGATGACCGCATCTTCCAGTACATATGCGCAGACATCCAG 121
QY 47 MetSerProArgValProGlnGlnAspTyrLeuSerGlnProProAlaArgValThrIle 66
Db 122 ATGAGTCCAGAGAGTCCCTCAGCAGAGACTGCTGTCTAAGCCCGCAGGATGATC 181
QY 67 LysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAspAspCysSer 86

```

```

|||||
Db 182 AAGATGAGTCAACCCCTAGCAGGTAAATGTTCCAGAACTCACTGATGAGTCACT 241
QY 87 ValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGlySer 106
Db 242 GTGAACAAGGTGGGAAGATGTGTGGCAGCCCGGATCTCTGTGGGATGAGTACGGCAGC 301
QY 107 TyrMetGluGlnLysHisIleProProProAsnMetThrThrAsnGlnLysArgValIle 126
Db 302 TACTATGAGAGAGAGATGTGTCCGCCCTCCCAATGTGACCAATATGAGCGGAGTATC 361
QY 127 ValProAlaAspProThrLeuTyrPheThrAspHisValArgGlnTyrPheGluTyrPala 146
Db 362 GTCCCTCAGATCTCTACTGTGTGAGCAGACACCATGTCGACAGTGGCTGAGTGGCGC 421
QY 147 ValLysGlnTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLys 166
Db 422 GTGAAGAATATGGCTCTCGATGTGAGCTCTTACTATTTCAGAATATCGATGGAGA 481
QY 167 GluLeuGlyLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAsp 186
Db 482 GAGCTGTGCAAGATGACAAAGATGACTTCCAGCGGCTCAGCGGACTACAAATGCCGAC 541
QY 187 IleLeuLeuSerHisLeuHisTyrLeuArgGlnThrProLeuProHisLeuThrSerAsp 206
Db 542 ATCTCTCTCACAATCTCCACTACTCAGAGAGACTCCCTCCACATCTGATCTCCGAT 601
QY 207 AspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAlaArgAsnThrGlyGly 226
Db 602 GACCTTATTAAGCTTTACAAACTCTCCAGGTTAATGATGCGCAGAAACACAGGGAGT 661
QY 227 AlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThr 246
Db 662 GCACCTTTATTTCCCAATACTTCACTAGTATATCCGAGGTAGCAGAAAGATTAACAAT 721
QY 247 ArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaThrThrSerHisSerHisPro 266
Db 722 AGGCCA-----
QY 267 ThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArg 286
Db 727 -----
QY 287 ProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGly 306
Db 728 -----GATCTTACCAATCTCTGGACCGACGACTACCCCTTGTCTAATCCAGGT 778
QY 307 SerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSer 326
Db 779 AGTGGCCAGATCCAGCTGTGGCAGTCTGCTCGCAACTCCTGTCAGACAGCTCCAACTCC 838
QY 327 AsnGlyIleThrThrGlnGlyThrAsnGlyGlyLysMetThrAspProAspGlnVal 346
Db 839 AACTGCATCTACCTGGGAGGACACCAAGGGAGTTCAAGATGACAGACCCGAGAGAGTGG 898
QY 347 AlaArgArgTyrGlyLysArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArg 366
Db 899 GCTGGCGCTGGGGGAGAGAGAGACCAAGCCCAACTGAACTATGACAACTAGCCGC 958
QY 367 AlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisProProGluSerSer 386
Db 959 GCCCTCCGTACTACTACGACAAACATCATGACCAAGTG-----
QY 387 MetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisGlyLysArgTyrAla 406
Db 1001 -----CACGGGAAGCGCTAGCC 1018
QY 407 TyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro-----
Db 1019 TACAGTTTACTTCCAGGGATTTGCCAGAGCCCTTGACGCCCACTCTGAGTGTCC 1078
QY 421 -----HisAlaHisProGlnLys 426

```

Db 1079 CTGTACAGTACCCCTCCGACCTGCATACATGGGCTCTATACAGCCGCCACCCAGAG 1138  
 Oy 427 MetasnpheValAlaProHisProAlaLeuProValThrSerSerPhePheAla 446  
 Db 1139 ATGACATTGTGTCTCCGCCACCTCCGCTCTCCAGACATCTTCAGATTCTTCT 1138  
 Oy 447 AlaProAsnProTyrTrpAsnSerProThrglyglylyIetyrProAsnThrArgLeuPro 466  
 Db 1199 TCCCGCAACCATCTGGAATTCAACGACCTGGGGGCACTCTACCCGAAACACTAGCTCCCA 1258  
 Oy 467 AlaAlaHisMetProSerHisLeuGlyThrTyrTrp 478  
 Db 1259 GCCAGCCATATGCCCTCTCACCTGGGCACTACTAC 1294  
 RESULT 10  
 ; Sequence 355, Application US/10144771  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig  
 ; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF  
 ; FILE REFERENCE: C1001321  
 ; CURRENT APPLICATION NUMBER: US/10/144,771  
 ; CURRENT FILING DATE: 2002-05-15  
 ; NUMBER OF SEQ ID NOS: 47235  
 ; SEQ ID NO 355  
 ; LENGTH: 2172  
 ; TYPE: DNA  
 ; ORGANISM: HUMAN  
 ; US-10-144-771-355  
 Alignment Scores:  
 Pred. No.: 3,77e-223 Length: 2172  
 Score: 2017.50 Matches: 384  
 Percent Similarity: 79.12% Conservative: 10  
 Best Local Similarity: 77.11% Mismatches: 7  
 Query Match: 77.96% Indels: 97  
 DB: 14 Gaps: 3  
 US-09-902-772-4 (1-478) x US-10-144-771-355 (1-2172)  
 Oy 1 MetAlaSerThrIleIleTysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 107 ATGGCCAGCACTATTAAAGAGCGCTGTGCACTTGAGCGAGGACAGCTCATTTTGTAG 166  
 Oy 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
 Db 167 TGTGCTTACGAAAGCCACACTGTGCTAGACAGAGATGACCCGATCTTCCAGTGAC 226  
 Oy 41 TysGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 Db 227 TATGGCCAGACATCCAAAGATGATGCCAGAGTCCCTCAGCAGGACTGGCTGTCCAAGCC 286  
 Oy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 287 CCAGCCAGGCTCACATCAGATGAGTGCACACCTTACAGTGAATGTTCCAGAGAC 346  
 Oy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
 Db 347 TCACCTGATGATGCTGACGTGAACAAGGTGGAGAGATGGGGGACCCGAGATACGTG 406  
 Oy 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120  
 Db 407 GGGGTGAGCTTACGGCCTCCTACATGAGAGAGAGCATGTGCCCTCCCAATATGATACACA 466  
 Oy 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140  
 Db 467 AATCAGGCGCAGAGATGCTGCTCCGAGATCTTCTGTGGAGCAACAGACATGTCCGA 526  
 Oy 141 GlnTrpLeuGlnTrpAlaValLysGlyTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 Db 527 CAGTGGCTGAGTGGGGGCGAAGAAATATGGCTCTCTCATGTGGAGCGCTTACTATTT 586  
 Oy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180

Db 587 CAGATATCGATGGAGAGAGAGCTGTGCAAGATGCAAAAGATGACTTCCAGCGGCTCAGC 646  
 Oy 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrIleLeuArgGluThrProLeu 200  
 Db 647 CCGAGCTTACATGCGCAGATCTTCTCTCACATCTCCAGTACCTCAGAGACATCCCTT 706  
 Oy 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220  
 Db 707 CCACATCTGACTTCCGATGATGAGCTTGATAGGCTTTACAAAACCTCCACGGTTAATCAT 766  
 Oy 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240  
 Db 767 GCCAGAAACACAGGGGGTGCAGCTTATATTTCCCAATACCTGATATATCCGAAGCT 826  
 Oy 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgSerAlaTrp 260  
 Db 827 ACGCAAGAAATTAACACTAGC----- 847  
 Oy 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280  
 Db 847 ----- 847  
 Oy 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300  
 Db 847 ----- 847  
 Oy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPglInPheLeuLeuLeuLeu 320  
 Db 848 -----CCAGTATGAGCCAGATCCAGCTGTGGGCGAGTCTGCTGCAACTCTG 895  
 Oy 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMet 340  
 Db 896 TCACACAGCTCCCACTCACTCATCATCAGCGGAGGACCAACAGGGGAGTTCAAGATG 955  
 Oy 341 ThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsn 360  
 Db 956 ACAGACCCGACGAGGTGGCTCGGGCTGGGGGAGAGAGAAAGACCAACCAATCATGAC 1015  
 Oy 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380  
 Db 1016 TATGACAAAGCTCAGCGCGGCTCGCTACTACTAGACCAAAAACATCATGACCAAGTG 1075  
 Oy 381 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400  
 Db 1075 ----- 1075  
 Oy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420  
 Db 1076 CACGGGAAGCGCTACGCTCAAGATTGACTTCCACGGGATGGCCAGGCTGACAGCC 1135  
 Oy 420 ----- 420  
 Db 1136 CACCTCTGAGTGTGCTCCGTACAACTACCCCTCCGACCTGCCATCATGAGGCTCTAT 1195  
 Oy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440  
 Db 1196 CACGCCACCCCGAGAGATGAATTTGTCTTCCACCTCCGCTCCGCTCCAGTACCA 1255  
 Oy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrArgLysIleTyr 460  
 Db 1256 TCTTCCAGTTCTTGTGCTCCCGAACCCATACTGGAATTCACCGATGGGGGCACTTAC 1315  
 Oy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTrp 478  
 Db 1316 CCGAACACTAGGCTCCAGCCAGCATATGCTCTCAGCTGGGCACTACTAC 1369  
 RESULT 11  
 ; PCT-US02-41414-1210  
 ; Sequence 1210, Application PC/TUS0241414  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Morris, David  
 ; APPLICANT: Engelhard, Eric



```

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71249/RMS/DCF
; CURRENT APPLICATION NUMBER: PCT/US02/41414
; CURRENT FILING DATE: 2002-12-26
; PRIOR APPLICATION NUMBER: US 09/7747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1210
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Homo sapiens
; PCT-US02-41414-1210

Alignment Scores:
Pred. No.: 3,82e-161 Length: 1359
Score: 1482.00 Matches: 291
Percent Similarity: 68.65% Conservative: 55
Best Local Similarity: 57.74% Mismatches: 80
Query Match: 57.26% Indels: 78
DB: Gaps: 9

US-09-902-772-4 (1-478) x PCT-US02-41414-1210 (1-1359)
Oy 1 MetLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 1 ATGGACGGGACTATTATGAGAGCTCTGTGGTGGAGCGAGCACCATCCCTCTTGAC 60
Oy 21 CysAlaTyrGlyLysSerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
Db 61 TCAGCGTACGGAGGGCGAGCCCATCTCCCAAGGCCACATGACGCTCGGGAGTCT 120
Oy 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGln 59
Db 121 GACTACGGGCGAGCCCGACCATGATCAACCCCTCCACAGCAGAGAGTGTGATCATCAG 180
Oy 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db 181 ---CCAGTGGAGGTCACGTCACAGCGGAGTAT-----GACCAACATGATGATCCAGG 231
Oy 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db 232 GAGTCTCCGGTGGACTGAGCGGTAGCAAAATGACGAAAGCTGTGGCGGAGCGGAGTCC 291
Oy 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---LleProProAsnMet 118
Db 292 AACCCCATGAACTACACAGCTATATGACGAGAAAGATGGCCCCCTCTCCACACATG 351
Oy 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTyrSerThrAspHis 138
Db 352 ACCACACAGAGAGAGAGTATGCTCCCGCAGACCCACACTGTGACAGAGAGAT 411
Oy 139 ValArgGlnTyrLeuGluTyrAlaValLysGluTyrGlyLeuProAspValAspLeu 158
Db 412 GTGAGGCAATGCTGAGTGGCCATAAAGATGACACTGTGAGAGATGACACATCC 471
Oy 159 LeuPheGlnAsnIleAspGlyLysGluLeuLysMetThrLysAspAspPheGlnArg 178
Db 472 TTTTTCGAGACATGATGATGACAGAACTGTGTAAAGAAAGAGAGACTTCTCCGC 531
Oy 179 LeuThrProSerTyrAsnAlaAspIleLeuSerHisLeuHisTyrLeuArgGlnThr 198
Db 532 GCCACACACCTCTACACACGGAAGTGTGTGTGCACACCTCAGTTACTTACGGGAAAGT 591
Oy 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
Db 592 TCACCTG----- 597
Oy 219 MethSAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
Db 598 -----CTGGCCTATATATACAACTCCACACC 624

```

```

Oy 239 GluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSer 258
Db 625 GACCAATCTTCACGATTAGTGTCAAAAGACCTTCTTATGACTCATGACAGAGAGA 684
Oy 259 AlaTyrThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
Db 685 GCTTGGGGCATTAAATCATGAAATCTTGGCTCAACAAAGATCTCTCCCTTGAGAGGGCACA 744
Oy 278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297
Db 745 ACGATCACTAAGATATACAGACCAACGGCCCGCAGATCCGTATACAGATCTCGGCCCG 804
Oy 298 ThrSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeu 317
Db 805 ACCAGCACTCCCTTACCCACACCTCGAAGCGGGCAGATCCAGCTGTGCAATTCCTCTTG 864
Oy 318 GluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrTyrGluGlyThrAsnGlyGlu 337
Db 865 GAGCTGCTCTCCGACAGCGCCACCGCATGTAATCATCTGAGAGGGAGACCAACGGGAG 924
Oy 338 PheLysMetThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysPro 357
Db 925 TTCAAAATGACGAGCCCGATGAGGTGGCCAGCGGCTGGGCGAGCCGGAAGAACGCC 984
Oy 358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377
Db 985 AACATGAAATTCAGACACACTAGCCGGCCCTCCGTTTATCATGATAAACATTTATG 1044
Oy 378 ThrLysValHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMet 397
Db 1045 ACCAAAGTG----- 1053
Oy 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417
Db 1054 -----CACGGCAAAAGATATGCTTACAAATTTGACTTCACAGGATTCGCCAGGCT 1104
Oy 418 LeuGlnPro----- 420
Db 1105 CTGCACGCCATCCGACCGAGTGTCTTACAAAGTACCTTCTGACATCTCCTACATG 1164
Oy 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
Db 1165 CCTTCCTACATGCGCCACACAGAGTGAAGTGTGCTCCCTCCCATCCATTCCTCATG 1224
Oy 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGly 457
Db 1225 CCTGTCACTTCTCCAGCTTCTTGGAGCCGATCACAATCTGAGACCTCCGCCACGGGG 1284
Oy 458 GlyIleTyrProAsn-----ThrArgLeuProAlaIleHisMetProSerHisLeu 474
Db 1285 GGATCTACCCCAACCCCAACAGTCCCGCATCTTACACCCACAGCTCCTTCACACTTA 1344
Oy 475 GlyThrTyrTyr 478
Db 1345 GGCAGCTACTAC 1356

RESULT 12
PCT-US02-41414-1209
; Sequence 1209, Application PC/TUS0241414
; GENERAL INFORMATION:
; APPLICANT: Engelhard, David
; APPLICANT: Morris, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71249/RMS/DCF
; CURRENT APPLICATION NUMBER: PCT/US02/41414
; CURRENT FILING DATE: 2002-12-26
; PRIOR APPLICATION NUMBER: US 09/7747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1613
; SOFTWARE: PatentIn version 3.1

```



ORGANISM: Homo sapiens  
US-10-342-887-710

## Alignment Scores:

Pred. No.: 1.55e-160 Length: 2957  
Score: 1482.00 Matches: 291  
Percent Similarity: 68.65% Conservative: 55  
Best Local Similarity: 57.74% Mismatches: 80  
Query Match: 57.26% Indels: 78  
DB: 14 Gaps: 9

US-09-902-772-4 (1-478) x US-10-342-887-710 (1-2957)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
DB 173 ATGACGCGGACTATTAAAGAGCCTCTGCGTGTGAGCAGCAGCAGCTCCCTTTGAC 232  
QY 21 CysAlaIArgly---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39  
DB 233 TCACGCTACGAGGAGCGGAGCCCATCTCCCAAGGCCACATGACTGCTCGGAGGACTCT 292  
QY 40 GluTyrGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGln 59  
DB 293 GACTACGAGGAGCCGCCACAGATACACCCCTCCACACGACGAGGTGATCAATCAG 352  
QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79  
DB 353 ---CCAGTACGAGGCTCAACGTCACAGCGGAGTAT-----GACCAATGATGATCCAGG 403  
QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99  
DB 404 GAGCTCCGCTGCTCAGGCTGAGCAATGACAGCAAGCGTGGCGGCGAGCGAGTCC 463  
QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---IleProProAsnMet 118  
DB 464 AACCCCATGACATCAACAGCTATATGAGAGAGATGAGCCCTCTCCCAACATG 523  
QY 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTyrSerThrAspHis 138  
DB 524 ACCACACGAGAGAGAGATCATGCTCCCGACAGCCACAGCTGTGACACAGAGCAT 583  
QY 139 ValArgGlnThrLeuGluTrrPalaValLysGluTyrGlyLeuProAspValAspIleLeu 158  
DB 584 GTGAGGCAATGGCTGAGTGGCCATTAAGAGTACGCTTGATGAGATCGACACATCC 643  
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178  
DB 644 TTTTCCAGAACATGATGGCAAGACTGTGTAATGACAGAGGAGACTTCTCCGC 703  
QY 179 LeuThrProSerTyrAsnAlaAspIleLeuSerHisLysTyrLeuArgGluThr 198  
DB 704 GCCACACCTCTACACAGAGAGTGTGTGTACACCTCATCTACAGGAGAGAGT 763  
QY 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218  
DB 764 TCACTG----- 769  
QY 219 MethIslaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238  
DB 770 -----CTGCCCTATTAATACAACTCCACACC 796  
QY 239 GluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSer 258  
DB 797 GACCAATCTCAGCATGATGAGTCAAGAGAGACCTCTTATGCTCAGTCAGAGAGGA 856  
QY 259 AlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277  
DB 857 GCTTGGGCAATAACATATTCGGCCTCAACAAAGCTCCCTCGTGGAGGGGACAA 916  
QY 278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297  
DB 917 ACATTCAGTAAAGATACAGAGCAAGGCCCAAGCCAGATCCGATGATCTGGGCCG 976

QY 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrrPglPheLeuLeu 317  
DB 977 ACCAGCAGTCCCTTACAGCAACCTCGAAGCGGAGAGATCCAGCTGTGCAATTCCTCTG 1036  
QY 318 GluLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrrPglGluGlyThrAsnGlyGlu 337  
DB 1037 GAGCTGCTCCGACAGCGCCACAGCGAGCTGTATCATCTGTGGAGGGAGCCAGCGGAG 1096  
QY 338 PheLysMetThrAspProAspGluValAlaArgArgTrrPglGlyLysArgLysSerLysPro 357  
DB 1097 TTCAAATGACGAGACCCGATGAGTGGCCAGCGCTGGGCGAGGAGGAGGAGGAGGAGCC 1156  
QY 358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTrrTrrTyrTyrAspLysAsnIleMet 377  
DB 1157 AACATGATATTCAGCAAGCTGAGCGGCGCTCCGTTATTCATGATATAAACAATTATG 1216  
QY 378 ThrLysValHisProProGluSerSerMetLysTrrTrrProSerAspLeuProTrrMet 397  
DB 1217 ACCAAAGTG----- 1225  
QY 398 SerSerTrrHisGlyLysArgTrrAlaTrrLysPheAspPheHisGlyIleAlaGlnAla 417  
DB 1226 -----CAGGCAAAAGATATGCTTACAAATTTGACTTCCACGGCATTTGCCAGGCT 1276  
QY 418 LeuGlnPro----- 420  
DB 1277 CTCGAGCGCAATCCGACGAGTGTGCTCATACAAAGTACACCTTCTGACATCTCCATATG 1336  
QY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437  
DB 1337 CCTTCTTACATGCCACGACAGCAAGAGTGAATCTTCTCCCTCCATTCATCTCCATG 1396  
QY 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTrrTrrProAsnSerProThrGly 457  
DB 1397 CCTGTCACTTCTCAGCTTCTTGGAGCGCATCAACAATGATGAGACTCCGCCACGAGG 1456  
QY 458 GlyIleTrrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474  
DB 1457 GGATTCACCCCAACCCCAACGATGCCCGGCATCTTAACACCACGAGGCTTACACTTA 1516  
QY 475 GlyThrTrrTrr 478  
DB 1517 GGCAGCTACATC 1528

RESULT 14  
PCT-US02-41414-1207  
Sequence 1207, Application PC/TUS0241414  
GENERAL INFORMATION:  
APPLICANT: Morris, David  
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER  
FILE REFERENCE: A-71249/RMS/DCF  
CURRENT APPLICATION NUMBER: PCT/US02/41414  
PRIOR FILING DATE: 2002-12-26  
PRIOR APPLICATION NUMBER: US 09/7747,377  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 09/7798,586  
NUMBER OF SEQ ID NOS: 1613  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1207  
LENGTH: 1359  
TYPE: DNA  
ORGANISM: Mus musculus  
PCT-US02-41414-1207

Alignment Scores:  
Pred. No.: 3.26e-160 Length: 1359  
Score: 1474.00 Matches: 291  
Percent Similarity: 68.25% Conservative: 53  
Best Local Similarity: 57.74% Mismatches: 82  
Query Match: 56.96% Indels: 78  
DB: 2 Gaps: 9





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2003, 04:15:54 ; Search time 2008.73 Seconds

(without alignments) 3853.895 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKKALSVSEDSQSLFE.....IYPNTRLPAAHNPMSHIGTY 478

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Command line parameters:  
-O/-cgn2.1/USPto.spool/US09902772/funat\_23072003\_093657\_14892/app\_query.fasta.1.1294  
-DB-EST-QFMT-fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST=45  
-DOCLALIGN=200 -THR.SCORE-pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptio -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09902772.ecgn.1.1.2619.6runat.23072003\_093657\_14892 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESOURCE -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=150 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1153	44.6	1045	13	BM456833
2	1141.5	44.1	880	14	BQ233264
3	1121	43.3	634	13	BM489636
4	1111.5	42.9	1089	14	BQ212127
5	1085	41.9	915	12	BG390291
6	1053	40.7	715	12	BG388025
7	1010	39.0	629	10	AM948986
8	992.5	38.4	928	12	BG255459
9	988	38.2	701	9	AU136709
10	984	38.0	642	13	BT1713036
11	984	38.0	668	10	BB660034
12	980	37.9	592	13	BM088063
13	968	37.4	827	9	AJ456498
14	948	36.6	675	10	BB639043
15	932	36.0	629	10	BB577887
16	882.5	34.1	1074	14	BQ953977
17	868.5	33.6	520	10	BE233460
18	845.5	32.7	949	13	BT558012
19	806	31.1	1015	9	AL554108
20	793	30.6	522	12	BF566705
21	769	29.7	516	12	BG691566
22	721	27.9	645	10	BB656030
23	715	27.6	666	12	BG390572
24	707	27.3	619	10	AM421310
25	701	27.1	671	13	BT588530
26	691	26.7	553	12	BG256864
27	686.5	26.5	1245	9	A1794504
28	666	25.7	1245	14	BQ212133
29	653.5	25.3	1068	14	BQ062565
30	648.5	25.1	872	14	BQ232274
31	643.5	24.9	740	9	AJ456506
32	639	24.7	894	14	BQ439561
33	619	23.9	815	9	AJ446143
34	616.5	23.8	756	9	AJ445352
35	609	23.5	787	10	BE559913
36	600.5	23.2	874	14	BQ215477
37	597.5	23.1	947	14	BQ642921
38	596.5	23.0	587	9	AL598619
39	588	22.7	651	10	BB442803
40	581.5	22.5	770	12	BG701574
41	580	22.4	788	9	AJ456689
42	573	22.1	907	10	BE561128
43	567.5	21.9	810	10	BE560516
44	558.5	21.6	972	12	BG257780
45	556	21.5	674	10	BB238163

#### ALIGNMENTS

RESULT 1  
LOCUS BM456833  
DEFINITION AGENCOURT\_6404041 NIH\_MGC\_92 Homo sapiens CDNA clone IMAGE:5583753  
ACCESSION BM456833  
VERSION BM456833.1 GI:18505873  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1045)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)





BASE COUNT 250 a 231 c 207 g 189 t 3 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4e-105 Length: 880  
 Score: 1141.50 Matches: 221  
 Percent Similarity: 95.02% Conservative: 8  
 Best Local Similarity: 91.70% Mismatches: 8  
 Query Match: 44.11% Indels: 4  
 DB: 14 Gaps: 1

US-09-902-772-4 (1-478) x BQ233264 (1-880)

QY 1 MetAlaSerThrIleuSerGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
 Db 74 ATGGCGGAGCTATTATTAAGGAGCGCTTATCAGTTGTGATGAGGACCGAGTCTGTGTGAG 133

QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGlu 40  
 Db 134 TGTGCTACGAGAGCGCCACACCTGGCTAAGACAGATGACCGCGCTCTCTCCACGAGAC 193

QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
 Db 194 TATGAGACAGACTTCCAGATGAGCCACGCGCTCCCTCAGAGATTGGCTGTCAACCC 253

QY 61 ProAlaIleValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 254 CCAGCCAGGAGTCCACCATCAAAATGAAATGTAACCTTACCGCAGGTGAAGGCTCAAGAAC 313

QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnVal 100  
 Db 314 TCTCTGTATGATGATCAGAGTGTGGCCAAAGGAGATGAGTGGGACGCGCCACAGACCGTT 373

QY 101 GlyMetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrThr 120  
 Db 374 GGGATGACACTACGCGACGCTACATGAGAGAGACATGCCACCCCAACATGACACGACG 433

QY 121 AsnGluIleValIleValIleValIleValIleValIleValIleValIleValIleVal 140  
 Db 434 AACGAGCGCAGAGTATGCTGTCGACGACATCTTACGCTATGATGATGACGACCATGTGGG 493

QY 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
 Db 494 CAGTGGCTGAGTGGGCGGTGAAGAATATGGCTTCCAGACGTCACATCTTGTATATC 553

QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 554 CAGAACATCGATGGAGAGAACTGTGCAAGATGACCAAGAGCGACTTCCAGAGGCTCAC 613

QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200  
 Db 614 CCCAGCTACACGCGGACATCTTCTCCACATCTCCACTACCTACGAGACATCTCTT 673

QY 201 ProHisLeuThrSerAspAspValAspLysAla-LeuGlnAsnSerProArg-LeuMetH 220  
 Db 674 CCACATTGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 733

QY 220 isAlaArgAsnThrGlyLysAlaThrPheIlePheProAsn---ThrSerValTyr 237  
 Db 734 ATGGCTAGAAACACAGAGGCGGTGTCAGCTTTTATTTTCCCAATACCTCAGTATAT 790

RESULT 3  
 LOCUS BM489636 634 bp mRNA linear EST 07-FEB-2002  
 DEFINITION pgm2n.pk011.124 Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cdna clone pgm2n.pk011.124 5' similar to sp|O90837|ERG.CHICK TRANSCRIPT. TRANSCRIPTIONAL REGULATOR ERG p1r1560754 transcription factor erg - chicken emb|CA54404.1| (x77159) ERG [Gallus gallus], mRNA sequence.  
 ACCESSION BM489636  
 VERSION BM489636.1 GI:18610567  
 KEYWORDS EST.

SOURCE chicken.  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianidae; Gallus.  
 REFERENCE 1 (bases 1 to 634)  
 AUTHORS Cogburn, L.A. and Monsonego-Ornan, E.  
 TITLE ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IRAS Animal Genome Project  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Larry A. Cogburn  
 University of Delaware  
 Townsend Hall, Newark, DE 19717, USA  
 Tel: 302-831-1335  
 Fax: 302-831-2822  
 Email: cogburnudel.edu, www.chickest.udel.edu.

FEATURES  
 source location/Qualifiers  
 1..634  
 /organism="Gallus gallus"  
 /strain="Commercial broiler and Ottawa Res. Centre strains 90 & 21"  
 /db\_xref="taxon:9031"  
 /clone="pgm2n.pk011.124"  
 /clone\_lib="Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n)"  
 /tissue\_type="Breast muscle, leg muscle and epiphyseal growth plate"  
 /dev\_stage="Breast, leg: Embryo(d19); post-hatch(d1,3,5,7,9,11 weeks); growth plate(d,7d,14d post-hatch)"  
 /lab\_host="E. coli EMDH10B"  
 /note="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each tissue (embryonic muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"

BASE COUNT 192 a 164 c 138 g 140 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.92e-103 Length: 634  
 Score: 1121.00 Matches: 210  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 43.32% Indels: 0  
 DB: 13 Gaps: 0

US-09-902-772-4 (1-478) x BM489636 (1-634)

QY 120 ThrAsnGluArgValIleValIleValIleValIleValIleValIleValIleVal 139  
 Db 3 ACCAATGACGAAAGATATGTCGACGACATCTTACGATGAGACGACGACATGTA 62

QY 140 ArgGlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeu 159  
 Db 63 CGGAGTGGCTGGAGTGGGCGAGTGAAGATGATGCTCCAGACGTGAGATCTTGTG 122

QY 160 PheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeu 179  
 Db 123 TTCAGAAACATTGATGGAAGAAGATTGTGTAATAATGACCAAGAATGACTTCCAGAGACT 182

QY 180 ThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrPro 199  
 Db 183 ACGCGGAGCTATTAACGAGATATCTCTCTCCACACTACCTACCTCAGAGACATCTT 242

QY 200 LeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMet 219  
 Db 243 CTTCCACATTTGACTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 302

QY 220 HisAlaArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGlu 239  
 Db 303 CATGCTAGAAACACAGAGGAGCGACCTTTATTTTCCCAATACATCATGTTACCGAGAA 362

QY 240 AlaThrGlnArgGlyIleThrArgProAspLeuProTyrGluGlnAlaArgArgSerAla 259  
 Db 363 GCAACGCAAAAGAAATACAAACAGGCCACATTTACTTATGAGCAAGGAGAGATCACAGC 422  
 QY 260 TrpThrSerHisSerHisProThrGlnSerIleValAlaThrGlnProSerSerSerThrVal 279  
 Db 423 TGGACGACTCACAGCATCCACCTAGTCAAAAGCTACCAACCATCATCTTCAACAGCTG 482  
 QY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299  
 Db 483 CCCAAACAGAAAGACACGACGTCCTCAGTTAGTCTTTCAGATTCTTGGACCGACACG 542  
 QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeu 319  
 Db 543 AGCGCTCTTGCATAATCCAGGAGATGGCCAGATACAGCATGACGATTCCTACTGAGCTT 602  
 QY 320 LeuSerAspSerSerAsnSerAsnCysIle 329  
 Db 603 CTGTGGACAGCTCCAACTCCAACTGCATC 632  
 RESULT 4  
 LOCUS BQ212127 1089 bp mRNA linear EST 02-MAY-2002  
 DEFINITION AGENCOURT\_7571287 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:6052136  
 ACCESSION BQ212127  
 VERSION BQ212127.1 GI:20392058  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1089)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rt@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
 http://image.lnl.gov  
 Plate: L14M13307 row: b column: 09  
 High quality sequence stop: 562.  
 Location/Qualifiers  
 1. 1089  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6052136"  
 /clone\_1id="NIH\_MGC\_92"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: testis; Vector: pCMV-SPORE6; Site: 1: NOTI;  
 Site\_2: SalI; Cloned unidirectionally; Oligo-dT primed.  
 Average insert size 2.5 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 292 a 314 c 257 g 225 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.07e-102 Length: 1089  
 Score: 1111.50 Matches: 229  
 Percent Similarity: 77.95% Conservative: 22  
 Best Local Similarity: 71.12% Mismatches: 46  
 Query Match: 42.95% Indels: 25  
 Db: 14 Gaps: 4  
 US-09-902-772-4 (1-478) x BQ212127 (1-1089)

QY 1 MetaLaserThrIleLysGluAlaLeuSerValSerGluAspGlnSerLeuPheGlu 20  
 Db 76 ATGGCCACACTATTAAAGAGCCCTTATCAGTTGTGATGAGACACAGTCTGTTTGGAG 135  
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetPheAlaSerSerSerSerGlu 40  
 Db 136 TGTGCTTACGGAAGCCACACCTGGCTTAAGACAGATGACCGGCTCTCTCCAGCCAC 195  
 QY 41 TyrGlyGlnThrSerIleLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60  
 Db 196 TATGACAGCATTTCCAAATATAGCCCAACGCTCCCTCNCAGATTTGGCTTCTCAACCC 255  
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
 Db 256 CCACGACGAGGTCACCATCAAAATGTAATGTAACCTGACCCAGTGGAATGGCTCAAGAAC 315  
 QY 81 SerProAspAspCysSerSerValAlaLysGlyValLysMetValSerSerSerAspVal 100  
 Db 316 TCTCCTGATGATGAGTGTGCGCCAAAGCGGGAAGATGGTGGCGACGCCACACCGCTT 375  
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThr 120  
 Db 376 GGGATGAACTACGGCAGCTACATGAGAGAGACACATGCCACCCCAACATGACACAGC 435  
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrPserThrAspHisValArg 140  
 Db 436 AACGACGGCAGAGTATCTGTCGCCAGAGATCTCAGCTAGTACAGACCAATGTGGCG 495  
 QY 141 GlnTrpLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160  
 Db 496 CAGTGGCTGGAGTGGCGCGTAAAGATATATGCGCTTCACAGCGTCACATCTTGTATTC 555  
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
 Db 556 CAGAAACATCCATGGGAAGGAACTGTGCAAGATGACCAACGACGCTTCCAGAGCTCAC 615  
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLe 200  
 Db 616 CCCAGCTACATGCGCCGACATCTTCTCTCACAATCTCCATCTCAGACGAGACTCTCT 675  
 QY 200 uProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMet 220  
 Db 676 TCCACATTTGACTACATGATGTTGNTAAAGCTTTCACAACTCTCCACGGGTAAATGC 735  
 QY 220 Is-AlaArgAsnThrGlyAlaThr-PheIlePheProAsnThr-----SerValTyr 237  
 Db 736 ATGGCTAAACCCAGGGGGGGCCAGCTTATTTTCCCAATCATTCAGAAATATTCCT 795  
 QY 237 rProGluAlaThrGlnArgIleThrArg-ProAspLeuProTyrGluGlnAlaArg 257  
 Db 796 GGAAGCTACCGCCAAAGATTTACACCTAGGGCCCAAAATTCACCTATYAGCCGCCCC 855  
 QY 257 rg-----SerAlaTrpThr-SerHisSerHisProThrGlnSerIleAlaThrGlnPro 274  
 Db 856 GGAAGATCACCCCTTGACCGCGGTGATGGCCACCC-----891  
 QY 275 SerSerSerThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIle 294  
 Db 892 -----CCGACGCCGAGAGTGGAAAGAGCTGGCCCCCCCAATTCCTCC----- 933  
 QY 295 LeuGlyProThrSerSerAspArgLeuAlaAsnProGlySerGlyGlnIleGlnLeu 312  
 Db 934 ---TCCCCGGGGGCCCAAAACCTGAAGAACACAGCGGTGCGCTTGAATAATC 984  
 RESULT 5  
 LOCUS BG390291 915 bp mRNA linear EST 12-MAR-2001  
 DEFINITION 602416255F1 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:4524852 5',  
 mRNA sequence.  
 ACCESSION BG390291  
 VERSION BG390291.1 GI:13283739  
 KEYWORDS EST.

SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLNL0429 row: m column: 13  
High quality sequence stop: 820.

FEATURES  
SOURCE  
Location/Qualifiers  
1..915  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4524852"  
/clone\_lib="NIH-MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: PCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC library."

BASE COUNT 246 a 252 c 227 g 190 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.27e-99 Length: 915  
Score: 1085.00 Matches: 238  
Percent Similarity: 77.26% Conservative: 10  
Best Local Similarity: 74.14% Mismatches: 26  
Query Match: 41.92% Indels: 51  
Gaps: 4

US-09-902-772-4 (1-478) x BG390291 (1-915)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
DB 59 ATGGCCACGACTATTAAAGAACCTTATCAGTTGTGAGTGAAGACCACTCGTTGTGAG 118  
QY 21 CysAlaIleTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40  
DB 119 TGTGCTTACGGAGACCCACACCTGGCTTAAGACAGATGACCGGCTCTCTCCACGCGAC 178  
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnPro 60  
DB 179 TATGACAGACTCTCAAGATGAGCGCCACGCGCTCAGAGATTGGCTGTCTCAACCC 238  
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
DB 239 CCAGCCAGGTCACCATCAAAATGGAATGAACCTACCAAGTCAATGGCTCAAGGAC 298  
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
DB 299 TCTCTGTATGAATGACAGTGTGGCCAAAGCGGGAAGATGTGGGACCCCAAGACCGGTT 358  
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120  
DB 359 GGGATGAACCTACGAGCTACATGGAGAGAGACCATGACCC -CCAACATGACCGACG 417  
QY 121 AsnGluArgValIleValProAla-AspProThrLeuThrPheThrAspHisValArg 140  
DB 418 AACGAGCGCAGATTATGCTCCACGACAGATCTCAACGCTATGAGATACAGACATGTGCG 477

QY 140 gGlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160  
DB 478 GCAGTGGCTGAGTGGGGGTAAAGAAATATGCTTCCAGACGTCACATCTTGTATT 537  
QY 160 eGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
DB 538 CCAGAACATCGATGGAGAGACTGTGCAGATGACCAAGAGACGACTTCCAGAGCTCAC 597  
QY 180 rProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuAlaGluThrPro-L 200  
DB 598 CCCAGCTACAAATGCCGACAT-CTTCTCTCATCATCTCCACTCCAGAGACATCTTC 656  
QY 200 euProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProAsnGluMetH 220  
DB 657 TTCCACATTGACTTCAAGTATGATGTGATTAAAGCTTACAAACTCTCCAGCGTTAAATC 716  
QY 220 lAlaIleArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGluA 240  
DB 717 ATGCTAGAAACACA----- 730  
QY 240 lArhGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaArg 260  
DB 731 -----GATTTACA-TATGAGCCCCCCAGAGATCAAGGC 763  
QY 260 rPThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValP 280  
DB 764 TGGAGGGTCACGG-CACCCACAGCCCAAGTCACATCTTCTTACAGTCCCAAA 822  
QY 280 rOlyThrGlu-----AspGlnArgProG 288  
DB 823 TGAACACAGAGTCGCTGATATGATGATCTTGGACAGTTCGCTGCAATGCGGGGCGGA 882  
QY 288 lNleuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 307  
DB 883 CTTTGATGCC-----GGGCGGGGAGC-----CCCGGCTCC 914

RESULT 6 715 bp mRNA linear EST 12-MAR-2001  
LOCUS BG388025 602412867P1 NIH-MGC\_92 Homo sapiens cDNA clone IMAGE:4521233 5',  
DEFINITION mRNA sequence.  
ACCESSION BG388025  
VERSION BG388025.1 GI:13281471  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLNL0420 row: f column: 18  
High quality sequence stop: 693.

FEATURES  
SOURCE  
Location/Qualifiers  
1..715  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4521233"  
/clone\_lib="NIH-MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: PCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed."

Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."

BASE COUNT 206 a 188 c 172 g 149 t

## ALIGNMENT SCORES:

Pred. No.: 2,74e-96 Length: 715  
Score: 1053.00 Matches: 195  
Percent Similarity: 97.58% Conservative: 7  
Best Local Similarity: 94.20% Mismatches: 5  
Query Match: 40.69% Indels: 0  
Gaps: 0

US-09-902-772-4 (1-478) x BG388025 (1-715)

QY 1 MetalaSerThrIleTylsGluAlaIleuSerValValSerGluAspGlnSerLeuPheGlu 20  
DB 94 ATGGCCAGCACTATTAAAGAACCTTATCACTTGTGAGTGAAGACCACTCTTTTGTAG 153  
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
DB 154 TGTGCTACGAGGAGCCACACCTGGCTAAGACAGAGATGACCGGCTCTCTCCACGAGC 213  
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPheSerGlnPro 60  
DB 214 TATGACAGACTTCCAGATGAGCCACGCTCCCTCAGAGATGGCTGTCTCAACCC 273  
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
DB 274 CCACCCAGGCGTACACATCAATGATGTAACCTTACCCAGGTGATGCTTCAAGAAC 333  
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
DB 334 TCTCTGATGATCAGTGTGGCCAAAGCGGGAAGATGTGGGACAGACACCGCT 393  
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120  
DB 394 GGGATGACCTAGCGCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 453  
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140  
DB 454 AACGAGCGCAGAGTATTCGTCAGAGATCTTACGGATGATGATGATGATGATGATG 513  
QY 141 GlnTyrPheGluTyrPheAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160  
DB 514 CAGTGGCTGGAGTGGGGGGAAGATATGCTTCCAGAGCAGCAACATCTGTATTC 573  
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
DB 574 CAGAACATCGATGGAG 633  
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200  
DB 634 CCGAGCTACAAAGCGGAGATCTCTCTCAATCTCAATATCTCAAGAGAGACCTCTT 693  
QY 201 ProHisLeuThrSerAspAsp 207  
DB 694 CCACATATGACTTCAGATGAT 714

RESULT 7  
AM948986 629 bp mRNA linear EST 31-MAY-2000  
LOCUS AM948986  
DEFINITION OVA-FR0005-110500-201-f02 FT0005 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AM948986  
VERSION AM948986.1 GI:8126760  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 629)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=422-QVA-FT0005-110  
500-201-f02&t3=2000-05-11&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 628.

## FEATURES

Location/Qualifiers  
1..629  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FR0005"  
/dev\_stage="Adult"

/note="Organ: prostate,tumor; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 166 a 173 c 171 g 119 t

## ORIGIN

## ALIGNMENT SCORES:

Pred. No.: 5.15e-92 Length: 629  
Score: 1010.00 Matches: 190  
Percent Similarity: 96.57% Conservative: 7  
Best Local Similarity: 93.14% Mismatches: 7  
Query Match: 39.03% Indels: 1  
Gaps: 0

US-09-902-772-4 (1-478) x AM948986 (1-629)

QY 6 LysGluAlaIleuSerValValSerGluAspGlnSerLeuPheGluCysAlaTyrGlySer 25  
DB 19 CAGGAACCTTATCAGTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 78  
QY 26 ProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyrGlyGlnThrSer 45  
DB 79 CCACACTGGCTTAAGACAGATATCCGCTCTCTCCACGACTTATGACAGACTTCC 138  
QY 46 LysMetSerProArgValProGlnGlnAspTyrPheSerGlnProProAlaArgValThr 65  
DB 139 AAGATGAGCCACGCGTCCCTCAGCAGAGATGGCTGTCTCAACCCACGAGGCTCAC 198  
QY 66 IleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAspAspCys 85  
DB 199 ATCAAAATGGAATTAACCTTACGAGAGTGAATGGCTCAAGGAAGTCTCTGTATGATGC 258  
QY 86 SerValAlaLysGlyLysMetValSerSerSerSerAspAsnValGlyMetAsnTyrGly 105  
DB 259 AGTGTGCCAAAGCGGAGAGATGGTGGAGCCACAGACCCGTTGGATGAACTACGGC 318  
QY 106 SerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsnGluArgVal 125  
DB 319 AGCTACATGAG 378

QY	1	MetLaserPrIleYSerGIuaLleuSerValValSerGIuaSpGIuSerLeuPhGlu	20
Db	88	ATGGCCAGCACATATTAAAGGAGCCCTTATTCAGATTGTGATGAGGACAGTGTGTG-TTTGAG	146
QY	21	CysAlaIyrGIYSerProHISleuAlaIylsthGILuMeTHrAlaSerSerSerGIu	40
Db	147	TGTGCTTACGGAAAGCCACACCTGTGCTTAAGACAGATGACCCGGCTCTCTCCAGGGAC	206
QY	41	TYIGLYGlnThSerILysMetSerProArgValProGlnGlnAspTrIleuSerGlnPro	60
Db	207	TATGACACGACTTCCAAAGATGAGCCCAAGCGGTCCCTCAGCAGATGTGGCTGTCAAGCC	266
QY	61	ProAlaIryalThrIleYSerMetGIuCysAnProAsnGlnValAsnGlySerArgAsn	80
Db	267	CCACCCAGAGGC-ACCATCAAAATGGAATGTAAACCTTGCCAGGTGAATGGCTCAAGGAC	325
QY	81	SerProAspAspCysSerValAlaIalysGILYlyMetValSerSerSerAspAsnVal	100
Db	326	TTCTCTTATGATATGACATGTGTGGCCAAAGCGGGAGATGGTGGGACCCAGACACCGTT	385
QY	101	GIYMetAsnTYrGIYSerTYrMetGIuGluYHsIleProProAsnMetLThrThr	120
Db	386	GGGATGAAGTACGGCAGCTACATGAGAGAGAACACATGCACCCCAACATGACCAGC	445
QY	121	AsnGIuArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHisValArg	140
Db	446	AACAGCGCAGAGTATTCGTGGCCAGCGAGATCCTACGTTAGAGATACAGCAATGTCCGG	505
QY	141	GlnTrpLeuGIuTrpAlaValILysGILTYrGIYLeuProAspValAspIleLeuPhe	160
Db	506	CAGTGGCTGAGTGGCGGTGACAGATATGGCTTCCAGACGTCAACATCTTGTATTC	565
QY	161	GlnAsnIleAspGILYlySGIuLeuCysILysMetThrLYsAspAspPheGlnArgLeuThr	180
Db	566	CAGAACATCGATGGAGGAAGAACTGTGCAAGTATGCCAACAGGACGACTTCCAGAGCTCAC	625
QY	181	ProSerTYrAsnAlaAspIleLeuLeuSerHisLeuHisTYrLeuArgGluThrProLeu	200
Db	626	CCAG--TACATAGCCGACATCCTTCTCTCATCTCCACTACCTCAAGAGAACCTCTCT	683
QY	201	ProHis-LeuThSerAspAspValAspLYsAlaLeuGlnAsnSerProArgLeuMetHI	220
Db	684	GCACATTTTGACTCAGATGATGTGTGATTAAGCTTACAAAATCTCCAGG-TTAAATGC	742
QY	220	SalArgAsnThrGIYlyAlaThrPheIlePheProAsnThSerValTYrProGlnAl	240
Db	743	TGCTAGAG---ACCCAGAGGGGTGTGTTTA-TTTCCAAATATATGATTT--TCTGAGC	795
QY	240	athrGlnArgIleThrThrArgProAspLeuProTYrGILGlnAlaArgSerAlaTr	260
Db	796	TACCAACAGATTA-----CCACTTGGCAGTCGAAAAACCGGTGATGCTGTCTG	843
QY	260	pThSerHisSerHisProThrGlnSerLYsAlaThcGlnProSerSerSerThrValPr	280
Db	844	GCACAGAC-----CAAGTCGTGATGCGGCCAACATCAAGAAAACAAGTGC	888
QY	280	o 280	
Db	889	C 889	
RESULT 9			
LOCUS	AUI36709	701 bp	mRNA linear EST 02-AUG-2002
DEFINITION	AUI36709 PLACE1 Homo sapiens cDNA clone PLACE1004911 5', mRNA		
ACCESSION	AUI36709		
VERSION	AUI36709.1	GI:10997248	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

**AUTHORS**  
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.

**TITLE**  
HRI human cDNA project

**JOURNAL**  
Unpublished (2000)

**COMMENT**  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomese@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

**FEATURES**  
Location/Qualifiers  
1..701  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="PLACE1004911"  
/clone\_lib="PLACE1"  
/tissue\_type="placenta"  
/note="Vector: pME18SFL3"  
BASE COUNT 199 a 185 c 166 g 145 t 6 others  
ORIGIN

**Alignment Scores:**  
Pred. No.: 1.03e-89 Length: 701  
Score: 988.00 Matches: 189  
Percent Similarity: 94.20% Conservative: 6  
Best Local Similarity: 91.30% Mismatches: 10  
Query Match: 38.18% Indels: 2  
DB: 9 Gaps: 0

US-09-902-772-4 (1-478) x A0136709 (1-701)

QY 1 MetAlaSerThrIleIleYsGluAlaLeuSerValSerGluAspGlnSerLeuPheGlu 20  
Db 85 ATGCCAGCAGCTATTAAAGAGACCTTATCACTGTGAGTGAGACACAGCTGTTGAG 144  
QY 21 CysAlaIleYrGlySerProHisLeuAlaLeuThrGluMetThrAlaSerSerSerGlu 40  
Db 145 TGTGCTTACGAGACGCCACCTGCTGAGACAGAGATGACCGGCTCTCTCCAGCCAC 204  
QY 41 TyrGlyGlnThrSerIleYsMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60  
Db 205 TATGACAGACTTCCAGATGAGCCACCGCTCCCTCAGCAGATGGCTGTCTCAACCC 264  
QY 61 ProAlaArgValThrIleIleYsMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
Db 265 CCACCCAGGCTCACATCAATGAATGTAACTGACCCAGGTGAATGGCTCAAGGAC 324  
QY 81 SerProAspAspCysSerValAlaIleYsGlyIleYsMetValSerSerSerAspAsnVal 100  
Db 325 TCTCTCTATGATCACTGCTGCGCAAGGCGGAGATGGTGGCAGCCAGACACCGCTT 384  
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluIleYsHisIleProProProAsnMetThrThr 120  
Db 385 GGGATGAACTACGCGCAGCTACATGAGAGAGACATGCCACCCCAACATATACACAG 444  
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140  
Db 445 AACGAGCGCAGAGTTATCTGTCAGCAGATCTACGCTATGATGATACAGACATGTCGG 504  
QY 141 GlnTrpLeuGluTrpAlaValIleYsGluIleYsMetValSerProAspValAspIleLeuPhe 160  
Db 505 CAGTGGCTGGAGTGGCGGTGAAGAAATGCGCTTCAGACGTCACATCTTGTATTC 564  
QY 161 GlnAsnIleAspGlyIleYsGluLeuCysIleYsMetThrTyrAspAspPheGlnArgLeuThr 180  
Db 565 CAGAACTCATGTGGAGANGAATCTGTCAAGATGACCAAGAGACGACTTTCNAGANGCTCAA 624

QY 180 pProSerTyrAsnAlaIleLeuLeuSerHisLeuHisTyrLeu-ArgGluThrProL 200  
Db 625 CCCAGCTACAGACCGGACATCTTCTGTACATCTCAGTACCTCAAGAGACTCTC 684  
QY 200 eupProHisLeuThrSer 205  
Db 685 TTTCACATTTGNCNTCA 701

**RESULT 10**  
B1713036  
LOCUS  
DEFINITION  
B1713036 642 bp mRNA linear EST 11-MAR-2002  
le01d08.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5085758 5'  
Similar to SW:ENG\_HUMAN P11308 TRANSCRIPTIONAL REGULATOR ERG ;,  
mRNA sequence.  
B1713036  
ACCESSION  
B1713036  
VERSION  
B1713036.1 GI:15688731  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 642)  
Mellon, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemishka, I., Secombe, M., Bressell, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Page, D., Wylie, T., Martin, D., Blistain, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas,  
M., Gibbons, M., McCann, R., Cole, R., Tsagarelashvili, R., Williams, T.,  
Jackson, Y. and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Other-ESTs: le01d08.x1  
Contact: Douglas Mellon, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmellon@molb.harvard.edu  
Laboratory was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center for information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@lm.wustl.edu)  
Seq primer: -40RP from Glbco  
High quality sequence stop: 473.  
Location/Qualifiers  
1..642  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5085758"  
/clone\_lib="HR85 islet"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site:1:  
NotI; Site:2: XhoI; cDNA made by oligo-dT priming.  
Size-selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@lm.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692.

**BASE COUNT** 178 a 171 c 163 g 129 t 1 others  
ORIGIN

**Alignment Scores:**  
Pred. No.: 2.3e-89 Length: 642  
Score: 984.00 Matches: 183  
Percent Similarity: 97.46% Conservative: 9  
Best Local Similarity: 92.89% Mismatches: 5  
Query Match: 38.02% Indels: 0  
DB: 13 Gaps: 0

RIKEN MouseESTs (Arakawa, T., et al. 2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho,  
Tel.: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>

```
QY 1 MetAlaSerThrIleArgLValAlaLeuSerValValSerGluuAspGlnSerLeuPheGlu 20
Db 80 ATGGCCAGACACTATTAGAGCGCCTTGTCAGTTGTGACGAGGACCACTCACTATTGGAG 135
QY 21 CysAlaTyrGlySerProHisLeuValAlaIleThrGluMetThrAlaSerSerSerGlu 40
Db 140 TGTGCTTACGAGGAACGGCACACTCTGGCTAAGACAGAGATGACCGCATCTCTTCACAGTAC 195
QY 41 TyrGlyGlnThrSerIleMetSerProAlaValProGlnGlnAspThrPLeuSerGlnPro 60
```

Db 200 TATGGCCAGACATCCAGATGATGCCAGACTCCCTCAGCAGACTGGCTGTCTCCAMCC 259  
Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
Db 260 CCAGCAGGAGGTCCACATCAATGATGAGTGCACACCTAGTACAGTGAATGTTCCAGAAC 319  
Qy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
Db 320 TCACCTGATGAGTGCAGTGTGACAAAGTGGAGAGATGGTGGCAGCCGAGATACGTGTG 379  
Qy 101 GlyMetAsnTrpGlySerTrpMetGluGlnLysHisIleProProAsnMetThrThr 120  
Db 380 GGGATGACCTACCGCAGCTACATGAGAGAGATGTCGCTCCCAATATATGCCACA 439  
Qy 121 AsnGlnArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140  
Db 440 AATGAGGCGCAGATGATGCTCCCTGAGATCTCTACTGTGTGAGACAGACCATGTCCGA 499  
Qy 141 GlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPhe 160  
Db 500 CAGTGGCTGAGTGGCGGCGTGAAGAATATGCGCTTCGATGTGGAGCTCTTACTATT 559  
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
Db 560 CAGATATCATGATGGGAAGAGAGCTGTGCAGATGACAAAGATGACTTCACAGCGCTCAG 619  
Qy 181 ProSerTrpAsnAlaAspIleLeuLeuSerHisLysLeuArg 196  
Db 620 CCGAGCTACATGCGCAGCATCTCTCTCATCATCTCNCACCTCAG 667

## RESULT 12

LOCUS BM088063 592 bp mRNA linear EST 19-NOV-2001  
DEFINITION 501436 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION BM088063  
VERSION BM088063.1 GI:16998691  
KEYWORDS EST.  
SOURCE ccr.  
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 592)  
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,  
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,  
Pettee,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
Keele,J.W.

## REFERENCE

## AUTHORS

## TITLE

Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)

JOURNAL MEDLINE  
COMMENT 21180013  
Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4356  
Fax: 402 762 4390

Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt. trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.

PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCAGTCAGCAGC  
Plate: 139 row: D column: 21  
Seq primer: ATTAGGTCAGCTATGAG.

## FEATURES

## SOURCE

1..592  
/organism="Bos taurus"  
/db\_xref="taxon:.9913"  
/clone\_id="MARC 2BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"

/note="Vector: pCMV SPOR6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from testis, thymus,  
semitendinosus muscle, longissimus muscle, pancreas,  
adrenal, and endometrium."  
BASE COUNT 152 a 177 c 159 g 104 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,17e-89 Length: 592  
Score: 980.00 Matches: 181  
Percent Similarity: 96.89% Conservative: 6  
Best Local Similarity: 93.78% Mismatches: 6  
Query Match: 37.87% Indels: 0  
Gaps: 0  
DB: 13  
US-09-902-772-4 (1-478) x BM088063 (1-592)

Qy 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20  
Db 12 ATGGCCAGCAGCTATTAAGAGCGTTATCACTCGTGAAGCAGGACAGTCACTGTTTGTAG 71  
Qy 21 CysAlaTrpGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40  
Db 72 TGCCCTTACGGCAGCCGCGACCTGCCAAGAGGACATGACCGGCTCTCTCCGGGAC 131  
Qy 41 TyrGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60  
Db 132 TATGACAGACATCCAGATGAGCCCGCTCCCTCAGCAGACTGGCTGTCTCAACCC 191  
Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80  
Db 192 CCAGCAGGAGGTCCACATCAATGATGAGTGAATGATGACCTTACAGTGAACGGCTCAAGAAC 251  
Qy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100  
Db 252 TCCCCTGATGATGCTCAGATGGCCAAAGGTGGGAAGATGGTGGGCGAGCCGAGACACCGTC 311  
Qy 101 GlyMetAsnTrpGlySerTrpMetGluGlnLysHisIleProProAsnMetThrThr 120  
Db 312 GGGATGAGTACACACAGCTACATGAGAGAGACATGACACCCCAACATGACACACC 371  
Qy 121 AsnGlnArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140  
Db 372 AACGAGCGCCCGCTCATCTGCTCCGACAGTCCACCTATGAGATACAGACATGTCGG 431  
Qy 141 GlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPhe 160  
Db 432 CAGTGGCTGAGTGGCGGAGTCAAGAATATGCGCTTCGAGCTTACATCTTATTTTC 491  
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180  
Db 492 CAGATATCATGAGGAGAGAGAGCTGTGCAGATGACCAAGATGACTTCAGAGGCTCACC 551  
Qy 181 ProSerTrpAsnAlaAspIleLeuLeuSerHisLysLeuArg 193  
Db 552 CCGAGCTACATGCGCAGCATCTCTCTCATCATCTCNCACCTCAG 590

## RESULT 13

LOCUS AJ456498 827 bp mRNA linear EST 22-APR-2002  
DEFINITION AJ456498 riken1 Gallus gallus cDNA clone 8ml372, mRNA sequence.  
ACCESSION AJ456498  
VERSION AJ456498.1 GI:20266594  
KEYWORDS EST.  
SOURCE chicken.  
ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 827)  
Buerstede,J.M.

## REFERENCE

## AUTHORS

## TITLE

JOURNAL Gallus gallus bursal lymphocyte EST  
Unpublished (2002)



## COMMENT

Contact: Buerstedde JM  
Cellular Immunology  
Heinrich-Pette-Institute  
Martinistr. 52, 20251 Hamburg, Germany  
Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.  
Location/Qualifiers

## FEATURES

## SOURCE

1. 827  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone\_1ib="r1ken1"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/note="CB inbred strain"

BASE COUNT 233 a 236 c 177 g 181 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1.41e-87 Length: 827  
Score: 968.00 Matches: 194  
Percent Similarity: 62.87% Conservative: 21  
Best Local Similarity: 56.73% Mismatches: 39  
Query Match: 37.40% Indels: 88  
Gaps: 5

US-09-902-772-4 (1-478) x AJ456498 (1-827)

OY 138 HtsValArgGlnTPLeuGlnTPAlaValLysGlnTyrGlyLeuProAspValAspIle 157  
Db 5 CACGTGGCCGCGTGGTGAATGGCCATTAAGGAGTGTGATTAAATGAGATGACACC 64  
OY 158 LeuLeuPheGlnAsnIleAspGlyLysGlnLeuLysGlnMetThrLysAspPheGln 177  
Db 65 ACCCTCTCCAGATATGATGCGCAAGAGCTCTGCCAAACGACAGATGACTTCTC 124  
OY 178 ArgLeuThrProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGlu 197  
Db 125 CGAACACCTCCCTCTACACACAGAAAGTCTGTGTCTACCTCAGTACTCAGGAA 184  
OY 198 ThrProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArg 217  
Db 185 AGT----- 187  
OY 218 LeuMetHisAlaArgAsnThrGlyValThrPheIlePheProAsnThrSerValTyr 237  
Db 188 -----AGCTACTGCTGGCTCTACAAATGCTCCTCCAC 220  
OY 238 ProGluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArg 257  
Db 221 ACAGAGCTTCCTCCTGCTTGCACCAAGAGTCTCTCTGTT----- 265  
OY 258 SerAlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSer 277  
Db 266 -----GCAGGAGCACCAAAAGCTGAAACAG 289  
OY 278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297  
Db 290 ACA-----ACAGAACACGACGCTCAGCCAGATCCCTTATCAATCTCGGCGCC 340  
OY 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuLeu 317  
Db 341 ACTAGTACTGCTTGGCAATCTCGGAGTGGGAGATACAACTATGACAGTCTCTCCTC 400  
OY 318 GluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLeu 337  
Db 401 GAGTGTCTGCTGGAGATGCCAGCCGCTGATACATGAGGAGGACCAATGGGGA 460  
OY 338 PheLysMetThrAspProAspGlnValAlaArgArgTrpGlyGluArgLysSerLysPro 357  
Db 461 TTCAAGATGACAGACCCAGATGAAAGTGCACGGCCTGGGAGAAACGAAACGACCC 520  
OY 358 AsnMetAsnTyrAspLysLeuSerSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377  
|||||

## RESULT 14

## BB639043

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

TITLE  
JOURNAL  
COMMENT

Db 521 AACATGATATGACAAAGCTGAGCCGACCTCTGATCTACTATGATAGAACATATATG 580  
OY 378 ThrLysValHisProProGlnLeuSerSerMetTyrLysTyrProSerAspLeuProTyrMet 397  
Db 561 ACCAAAGT----- 589  
OY 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417  
Db 590 -----CATGGCAAAAGGATGCTTACAAATTTGACTTTCATGAGCATTTGCCAGGCT 640  
OY 418 LeuGlnPro----- 420  
Db 641 CTCACAGCTCATCCACATGATATCATATGATCAAGATATCATGATGATCTCTCATATG 700  
OY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437  
Db 701 CCTTCTTACCATGCGCCACGACAGAGAGTGAACCTTGTACCCACACCTCTCTCTATG 760  
OY 438 ProValThrSerSerSerPhePheAlaProAsnProTyrTrpAsnSerProThrGly 457  
Db 761 CCGTCAATATCATTCAGTTTCTTGGAGACAGCTCACCCTTATTTGACCTCCCTGCTGGGA 820  
OY 458 GlyIle 459  
Db 821 AGCAT 826

BB639043 675 bp mRNA linear EST 31-AUG-2001  
BB639043 RIKEN full-length enriched, 3 days neonate thymus Mus  
musculus cDNA clone A630068021 5', mRNA sequence.  
BB639043  
BB639043.1 GI:15401323  
EST.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 675)  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,  
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,  
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
Tagami, M., Tagawa, A., Takahashi, F., Takeeda, Y., Tanaka, T., Toya, T.,  
Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
Unpublished (2001)  
Contact: Yoshinide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
url: <http://genome.gsc.riken.go.jp/>  
Carninci, P., Shibata, Y., Hayatsu, M., Sugahara, Y., Shibata, K., Itoh,  
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
Wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E.,  
Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,  
Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a





**THIS PAGE BLANK (USPTO)**